

# QPTAS for Geometric Set-Cover Problems via Optimal Separators

Nabil H. Mustafa  
 Université Paris-Est,  
 Laboratoire d'Informatique Gaspard-Monge,  
 Equipe A3SI, ESIEE Paris.  
 mustafan@esiee.fr

Rajiv Raman  
 Dept. of Computer Science.  
 IIIT, Delhi.  
 rajiv@iiitd.ac.in

Saurabh Ray  
 Computer Science, New York University, Abu Dhabi.  
 saurabh.ray@nyu.edu

## Abstract

Weighted geometric set-cover problems arise naturally in several geometric and non-geometric settings (e.g. the breakthrough of Bansal-Pruhs (FOCS 2010) reduces a wide class of machine scheduling problems to weighted geometric set-cover). More than two decades of research has succeeded in settling the  $(1 + \epsilon)$ -approximability status for most geometric set-cover problems, except for four basic scenarios which are still lacking. One is that of weighted disks in the plane for which, after a series of papers, Varadarajan (STOC 2010) presented a clever *quasi-sampling* technique, which together with improvements by Chan *et al.* (SODA 2012), yielded a  $O(1)$ -approximation algorithm. Even for the unweighted case, a PTAS for a fundamental class of objects called pseudodisks (which includes disks, unit-height rectangles, translates of convex sets etc.) is currently unknown. Another fundamental case is weighted halfspaces in  $\mathbb{R}^3$ , for which a PTAS is currently lacking. In this paper, we present a QPTAS for all of these remaining problems. Our results are based on the separator framework of Adamaszek-Wiese (FOCS 2013, SODA 2014), who recently obtained a QPTAS for weighted independent set of polygonal regions. This rules out the possibility that these problems are APX-hard, assuming  $\mathbf{NP} \not\subseteq \mathbf{DTIME}(2^{\text{polylog}(n)})$ . Together with the recent work of Chan-Grant (CGTA 2014), this settles the APX-hardness status for all natural geometric set-cover problems.

# 1 Introduction

One of the fundamental optimization problem is the set-cover problem: given a range space  $(X, \mathcal{R})$  consisting of a set  $X$  and a set  $\mathcal{R}$  of subsets of  $X$  called the *ranges*, the objective is to compute a minimum-sized subset of  $\mathcal{R}$  that covers all the points of  $X$ . Unfortunately in the general case, it is strongly NP-hard; worse, it is NP-hard to approximate the minimum set-cover within a factor of  $c \log n$  of the optimal [31] for some constant  $c$ .

A natural extensively-studied occurrence of the set-cover problem is when the ranges are derived from geometric objects. For example, given a set  $P$  of  $n$  points in the plane and a set  $\mathcal{R}$  of disks, the set-cover problem for disks asks to compute a minimum cardinality subset of disks whose union covers all the points of  $P$ . Unfortunately, computing the minimum cardinality set-cover remains NP-hard even for basic geometric objects, such as unit disks in the plane. Effort has therefore been devoted to devising approximation algorithms for geometric set-cover problems (see [7, 29, 15, 12, 3, 11, 14, 22, 19, 29, 8, 17, 16, 9, 28] for a few examples). Nearly all the effort has been for the following natural and fundamental categories of geometric objects: *halfspaces*, *balls* (and generally, *pseudodisks*), *axis-parallel rectangles*, *triangles* and *objects parameterized by their union-complexity* (a set of regions  $\mathcal{R}$  has union complexity  $\phi(\cdot)$  if the boundary of the union of any  $r$  of the regions has at most  $r\phi(r)$  intersection points). An important version is the *weighted* setting, where one seeks to find the minimum-weight set-cover (e.g., see the breakthrough of Bansal-Pruhs [6] who reduced a broad class of machine scheduling problems to weighted set-cover problems).

Research during the past three decades has, in fact, been able to largely answer the question of the existence of a PTAS, or provability of APX-hardness for these problems for the uniform case, where one is minimizing the cardinality of the set-cover. For the more general weighted case, there has been considerable progress recently – a  $O(\log \phi(\text{OPT}))$ -approximation as a function of the union-complexity is possible via the quasi-uniform sampling technique of Varadarajan [33] and its improvement by Chan *et al.* [9]. On the other hand, recently Chan-Grant [8] proved APX-hardness results for the set-cover problem for a large class of geometric objects. We also point out (Appendix C) that for any integer  $s$ , there exist  $O(s)$ -sided polygons with union complexity  $n2^{\alpha(n)^{O(s)}}$  for which set-cover is inapproximable within  $\Omega(\log s)$ . Also, as any set-system with sets of size at most  $s$  can be realized by halfspaces in  $\mathbb{R}^{2s}$ , a  $\Omega(\log d)$  lower-bound follows for approximability of halfspaces set-cover in  $\mathbb{R}^d$  (Appendix C). This lower-bound requires  $d \geq 4$ , leaving open the interesting question of approximation schemes for weighted halfspaces in  $\mathbb{R}^3$ .

See the table for the current status of geometric set-cover. The four open cases present a challenge as the current state-of-the-art methods hit some basic obstacles: the approximation algorithms for weighted halfspaces, balls and pseudodisks use LP-rounding with  $\epsilon$ -nets [33, 9], and so provably cannot give better than  $O(1)$ -approximation algorithms. LP rounding was avoided by the use of local-search technique [28] to give a PTAS for halfspaces ( $\mathbb{R}^3$ ) and disks ( $\mathbb{R}^2$ ); however *i*) for fundamental reasons it is currently limited to the unweighted case, and *ii*) does not extend to pseudodisks.

Object	Uniform	Weighted
Halfspaces ( $\mathbb{R}^2$ )	Exact	Exact
Halfspaces ( $\mathbb{R}^3$ )	PTAS	?
Halfspaces ( $\mathbb{R}^4$ )	APX-H	APX-H
Balls ( $\mathbb{R}^2$ )	PTAS	?
Pseudodisks ( $\mathbb{R}^2$ )	?	?
Balls ( $\mathbb{R}^3$ )	APX-H	APX-H
A-P Rects ( $\mathbb{R}^2$ )	APX-H	APX-H
A-P Rects ( $\mathbb{R}^d$ )	APX-H	APX-H
Triangles ( $\mathbb{R}^2$ )	APX-H	APX-H
Linear U-C ( $\mathbb{R}^2$ )	APX-H	APX-H

In this paper, we make progress on the approximability status of the remaining four open cases by presenting a quasi-polynomial time approximation scheme (QP-TAS) for all these problems. This rules out the possibility that these problems are APX-hard, assuming  $\text{NP} \not\subseteq \text{DTIME}(2^{\text{polylog}(n)})$ . Together with the previous work showing hardness results or PTAS, this settles

the APX-hardness status for all natural geometric set-cover problems.

The motivation of our work is the recent progress on approximation algorithms for another fundamental geometric optimization problem, maximum independent sets in the intersection graphs of geometric objects, where  $(1 + \epsilon)$ -approximation algorithms (or even constant factor approximation algorithms) are not known for many objects. In a recent breakthrough, Adamaszek-Wiese [1, 2] presented a QPTAS for computing weighted maximum independent set for a variety of geometric objects (e.g., axis-parallel rectangles, line-segments, polygons with polylogarithmically many sides) in the plane (the algorithm runs in time  $2^{\text{poly}(\log n/\epsilon)}$ ). We now sketch their main idea for approximating the maximum independent set for weighted line-segments in the plane, for which let  $\text{OPT}$  be the optimal solution. The key tool is the existence of a closed polygonal curve  $\mathcal{C}$  (with few vertices) that intersects segments in  $\text{OPT}$  with small total weight, and at least a constant fraction of the total weight of  $\text{OPT}$  lies in the two regions created by  $\mathcal{C}$ . Hence one can guess the curve  $\mathcal{C}$  (which does not require knowing  $\text{OPT}$ )<sup>1</sup>, and then return the union of solution of the two sub-problems (which are solved recursively). With appropriate parameters, the loss incurred by throwing away the segments intersecting  $\mathcal{C}$  is at most  $\epsilon$ -th fraction of the optimal solution, yielding a  $(1 - \epsilon)$ -approximation in quasi-polynomial time.

Let us consider how the above technique can be made to work for the set cover problem. Assume that we are given a set  $P$  of  $n$  points and a set  $\mathcal{R}$  of  $m$  weighted disks and our goal is to pick a minimum-weight set-cover from  $\mathcal{R}$ . We can again consider the optimal solution  $\text{OPT}$  and hope to find a curve  $\mathcal{C}$  which intersects objects in  $\text{OPT}$  with small total weight and has a constant fraction of the weight of  $\text{OPT}$  in the interior as well as the exterior. However such a curve does not always exist – consider, e.g., a case where the optimal solution consists of a set of disks that share a common point (not necessarily in  $P$ ). Crucially, unlike the independent set problem, the objects in the optimal set-cover are not disjoint. This dooms any separator-based approaches for the set-cover problem.

Surprisingly, we show that nevertheless there still exists a curve  $\mathcal{C}$  (which may, in fact, intersect all the disks in  $\mathcal{R}$ !) such that solving the induced sub-problems in the interior and exterior of  $\mathcal{C}$  and combining them leads to a near-optimal solution (Theorem 3.1). The problem is further complicated by arbitrary weights on the disks. As a result, several promising approaches (including the quasi-sampling technique of Varadarajan [33]) fail. Fortunately, generalizing the problem to pseudodisks and then using structural properties of pseudodisks melded with randomized ordering and probabilistic re-sampling techniques works out.

Such separator based techniques do not work in three dimensions (even for the independent set problem, one can show that there exists a set of disjoint segments in  $\mathbb{R}^3$  so that there is no compactly-represented polyhedral separator). In fact, even for unit balls in  $\mathbb{R}^3$  all containing a common point, the set cover problem is APX-hard [8]! However, when the objects are halfspaces in  $\mathbb{R}^3$ , we prove the existence of a polyhedral separator that allows us to get a QPTAS. This shows that the set-cover problem for halfspaces is the only natural problem in  $\mathbb{R}^3$  that is not APX-hard.

## 2 Preliminaries

Let  $\mathcal{R} = \{R_1, \dots, R_n\}$  be a set of weighted  $\alpha$ -simple regions in the plane, where a bounded and connected region in the plane is called  $\alpha$ -simple if its boundary can be decomposed into at most  $\alpha$   $x$ -monotone arcs<sup>2</sup>.

<sup>1</sup>The guessing is actually done by enumerating all possible curves. The fact that  $\mathcal{C}$  has a small number of vertices allows efficient enumeration.

<sup>2</sup>Note that whether a region is  $\alpha$ -simple depends on choice of axes.

For any  $\alpha$ -simple region  $R$ , we denote by  $\Gamma(R)$  a set of at most  $\alpha$   $x$ -monotone curves that its boundary can be decomposed into. For a set  $\mathcal{R}$  of  $\alpha$ -simple regions, we define  $\Gamma(\mathcal{R})$  to be the set  $\bigcup_{R \in \mathcal{R}} \Gamma(R)$ . Let  $w_i$  denote the weight of the region  $R_i \in \mathcal{R}$ , and  $w(\mathcal{S})$  be the total weight of the regions in  $\mathcal{S}$  (set  $W = w(\mathcal{R})$ ). The regions in  $\mathcal{R}$  need not be disjoint<sup>3</sup>. A collection of compact simply connected regions in the plane is said to form a family of *pseudodisks* if the boundaries of any two of the regions intersect at most twice. The union complexity of a set of pseudodisks is linear [30]. For technical reasons we will assume that pseudodisks in this paper are  $\alpha$ -simple for some constant  $\alpha$ . This restriction is not crucial, and can be removed [20]. A collection of regions  $\mathcal{R}$  is said to be *cover-free* if no region  $R \in \mathcal{R}$  is covered by the union of the regions in  $\mathcal{R} \setminus R$ . For any closed Jordan curve  $\mathcal{C}$ , we denote the closed region bounded by it as *interior*( $\mathcal{C}$ ) and the closed unbounded region defined by it as *exterior*( $\mathcal{C}$ ). Given  $\mathcal{R}$ , we denote by  $\mathcal{R}_{in}(\mathcal{C})$  ( $\mathcal{R}_{ext}(\mathcal{C})$ ) the subset of the regions that lie in *interior*( $\mathcal{C}$ ) (*exterior*( $\mathcal{C}$ )). Similarly if  $P$  is a set of points, we denote by  $P_{in}(\mathcal{C})$  ( $P_{ext}(\mathcal{C})$ ) the subset of points lying in *interior*( $\mathcal{C}$ ) (*exterior*( $\mathcal{C}$ )).

**VC-dimension and  $\epsilon$ -nets [24].** Given a range space  $(X, \mathcal{F})$ , a set  $X' \subseteq X$  is *shattered* if every subset of  $X'$  can be obtained by intersecting  $X'$  with a member of the family  $\mathcal{F}$ . The VC-dimension of  $(X, \mathcal{F})$  is the size of the largest set that can be shattered. Given a set system  $(X, \mathcal{F})$  where each element of  $X$  has a positive weight associated with it, and a parameter  $0 < \epsilon < 1$ , an  $\epsilon$ -net is a subset  $Y \subseteq X$  s.t. for any  $F \in \mathcal{F}$  with weight at least an  $\epsilon$  fraction of the total weight,  $Y \cap F \neq \emptyset$ . The  $\epsilon$ -net theorem (Haussler-Welzl [21]) states that there exists an  $\epsilon$ -net of size  $O(d/\epsilon \log 1/\epsilon)$  for any range space with VC-dimension  $d$ .

**QPT-partitionable problems.** Given an optimization problem  $\mathcal{O}$ , let  $\text{OPT}_{\mathcal{O}}(I)$  denote the optimal solution of  $\mathcal{O}$  on the instance  $I$ , and let  $w(\text{OPT}_{\mathcal{O}}(I))$  be the weight of this optimal solution. We assume  $\mathcal{O}$  is a minimization problem; similar statements hold for the maximization case.

**Definition 2.1.** A problem  $\mathcal{O}$  is quasi-polynomial time partitionable (QPT-partitionable) if, given any input  $I$  and a parameter  $\delta > 0$ , there exist a constant  $c < 1$ ,  $k = O\left(n^{(\log n/\delta)^{O(1)}}\right)$ , and instance pairs  $(I_l^1, I_r^1), \dots, (I_l^k, I_r^k)$  (computable in time polynomial in  $k$ ), and an index  $j$ ,  $1 \leq j \leq k$ , such that i)  $\max\{w(\text{OPT}_{\mathcal{O}}(I_l^j)), w(\text{OPT}_{\mathcal{O}}(I_r^j))\} \leq c \cdot w(\text{OPT}_{\mathcal{O}}(I))$ , ii)  $\text{OPT}_{\mathcal{O}}(I_l^j) \cup \text{OPT}_{\mathcal{O}}(I_r^j)$  is a feasible solution, and iii)  $w(\text{OPT}_{\mathcal{O}}(I_l^j)) + w(\text{OPT}_{\mathcal{O}}(I_r^j)) \leq (1 + \delta)w(\text{OPT}_{\mathcal{O}}(I))$ .

The next lemma follows immediately from recursive divide-and-conquer:

**Lemma 2.1.** If a problem  $\mathcal{O}$  is QPT-partitionable, and if for any instance  $I$ ,  $w(\text{OPT}_{\mathcal{O}}(I)) \geq 1$ , then one can compute a  $(1 + \epsilon)$ -approximate solution for  $\mathcal{O}$  in time  $O\left(n^{(\frac{1}{\epsilon} \cdot \log w(\text{OPT}_{\mathcal{O}}(I)) \cdot \log n)^{O(1)}}\right)$ .

*Proof.* The algorithm will return an approximate solution  $\text{APPROX}_{\mathcal{O}}(I)$  as follows. Let  $T = w(\text{OPT}_{\mathcal{O}}(I))$ , and set  $\delta = \Theta(\epsilon/\log T)$ . Construct the  $k$  instance pairs  $(I_l^1, I_r^1), \dots, (I_l^k, I_r^k)$ , where  $k = O\left(n^{(1/\epsilon \cdot \log T \cdot \log n)^{O(1)}}\right)$ . For each  $i = 1 \dots k$ , compute  $\text{APPROX}_{\mathcal{O}}(I_l^i)$  and  $\text{APPROX}_{\mathcal{O}}(I_r^i)$  recursively and return the solution  $\text{APPROX}_{\mathcal{O}}(I_l^j) \cup \text{APPROX}_{\mathcal{O}}(I_r^j)$ , where  $j = \arg \min_i w(\text{APPROX}_{\mathcal{O}}(I_l^i)) + w(\text{APPROX}_{\mathcal{O}}(I_r^i))$ . We can prune the recursion tree at the level  $l = O(\log T)$  since for the right choice of  $i$  at each recursion, the weight of the optimal solution falls by a constant factor with every recursive call. The size of the tree is at most  $(2k)^l = O\left(n^{(1/\epsilon \cdot \log T \cdot \log n)^{O(1)}}\right)$ . It can be shown inductively that the approximation factor of a sub-problem  $t$  levels away from the lowest level is  $(1 + \delta)^t$ . Thus the approximation factor at the root is  $(1 + \delta)^l \leq (1 + \epsilon)$ , with appropriate constants in the definition of  $\delta$ . The time taken by the algorithm is  $O\left(n^{(1/\epsilon \cdot \log T \cdot \log n)^{O(1)}}\right)$ .  $\square$

<sup>3</sup>We assume  $\mathcal{R}$  to be in general position, so no three regions boundaries intersect at the same point.

**Geometric separators.** A  $\delta$ -separator for  $\mathcal{R}$ , given  $\delta > 0$ , is a simple closed curve  $\mathcal{C}$  in the plane such that the number of regions of  $\mathcal{R}$  completely inside (and outside)  $\mathcal{C}$  is at most  $2W/3$  (such a curve is called *balanced*), and the total weight of the regions in  $\mathcal{R}$  intersecting  $\mathcal{C}$  is at most  $\delta W$ . The goal is to show the existence, given  $\mathcal{R}$  and  $\delta > 0$ , of separators of small combinatorial complexity as a function of  $n$ ,  $m$  (number of intersections in  $\mathcal{R}$ ),  $\alpha$  and  $\delta$ . The existence of small  $\delta$ -separators was the core of the result of [1, 2]; later several authors noted [27, 20] that the construction in [1] can be made optimal using the techniques of constructing cuttings and  $\epsilon$ -nets (i.e, the probabilistic re-sampling technique) [10, 24, 12, 5].

We state two separator results that we will be using in our algorithm. For completeness we outline their proof in Appendix D.

**Theorem 2.2** ([27, 20]). *Given a set  $\mathcal{R}$  of  $n$  weighted  $\alpha$ -simple regions (with total weight  $W$ , and no curve having weight more than  $W/3$ ) with disjoint interiors, and a parameter  $\delta > 0$ , there exists a simple closed curve  $\mathcal{C}$  such that i) the total weight of the regions intersecting  $\mathcal{C}$  is at most  $\delta W$ , and ii) the total weight of the regions completely inside or outside  $\mathcal{C}$  is at most  $2W/3$ . Furthermore the complexity of  $\mathcal{C}$  is  $T = O(\alpha/\delta)$ . That is  $\mathcal{C}$  can be completely described by a sequence of at most  $T$  curves of  $\Gamma(\mathcal{R})$  and additional at most  $T$  bits. Furthermore this is optimal; even when  $\mathcal{R}$  is set of disjoint line segments, any  $\mathcal{C}$  satisfying these two properties must have  $\Omega(1/\delta)$  bends.*

In the case when the regions have uniform weights (say each region has weight one) but are not necessarily disjoint:

**Theorem 2.3** ([27]). *Given a set  $\mathcal{R}$  of  $n$   $\alpha$ -simple regions in the plane with  $m$  intersections, and a parameter  $r$ , there exists a simple closed curve  $\mathcal{C}$  such that i) the number of regions in  $\mathcal{R}$  intersecting  $\mathcal{C}$  are  $O((m + \frac{\alpha^2 n^2}{r})^{1/2})$ , and ii) the total number of regions completely inside or outside  $\mathcal{C}$  is at most  $2n/3$ . Furthermore, complexity of  $\mathcal{C}$  is  $T = O((r + \frac{mr^2}{\alpha^2 n^2})^{1/2})$ . That is,  $\mathcal{C}$  can be completely described by a sequence of at most  $T$  curves of  $\Gamma(\mathcal{R})$  and at most  $T$  additional bits.*

The technical condition that the regions are  $\alpha$ -simple in the theorems above can be removed [20].

**Geometric Set Cover.** Let  $\mathcal{R} = \{R_1, \dots, R_n\}$  be a set of weighted regions (in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ ) and let  $P$  be a finite set of points in the plane. The goal is to compute a subset  $\mathcal{Q} \subseteq \mathcal{R}$  minimizing the total weight  $w(\mathcal{Q})$  so that  $P \subseteq \cup_{Q \in \mathcal{Q}} Q$ . We will denote an optimal solution  $\mathcal{Q}$  for an instance of the problem given by a set of regions  $\mathcal{R}$  and a set of points  $P$  by  $\text{OPT}(\mathcal{R}, P)$ , and its weight by  $w(\text{OPT}(\mathcal{R}, P))$ .

**Claim 2.4.** *If there exists a QPTAS for set-systems  $(\mathcal{R}', P')$  where i) each  $R \in \mathcal{R}'$  has weight  $w(R) \geq 1$ , and ii) the weight of the optimal set-cover for  $(\mathcal{R}', P')$  is  $O(n/\epsilon)$ , then there exists a QPTAS for the minimum-weight set cover for a set-system  $(\mathcal{R}, P)$  with arbitrary weights.*

*Proof.* Let  $\mathcal{Q}$  be a minimum-weight set-cover for  $(\mathcal{R}, P)$ . First guess the maximum weight region in  $\mathcal{Q}$ , say of weight  $w_{\max}$  (there are  $n$  such choices). Then by exponential search on the interval  $[w_{\max}, nw_{\max}]$ , one can guess the weight of  $\mathcal{Q}$  within a  $(1 + \epsilon/3)$  factor (there are  $O(\log_{1+\epsilon} n)$  such choices). Let  $w_{\text{appr}}$  be this weight, satisfying  $w(\mathcal{Q}) \leq w_{\text{appr}} \leq (1 + \epsilon/3)w(\mathcal{Q})$ . Set  $\mathcal{R}' \subset \mathcal{R}$  to be the set of regions with weight at least  $\epsilon w_{\text{appr}}/n$ , and  $\mathcal{R}'' = \mathcal{R} \setminus \mathcal{R}'$ . Let  $P' \subseteq P$  be the set of points not covered by  $\mathcal{R}''$ , and construct a  $(1 + \epsilon/3)$  approximate set-cover  $\mathcal{Q}'$  to  $(\mathcal{R}', P')$ . Return  $\mathcal{Q}' \cup \mathcal{R}''$  as a set-cover for  $(\mathcal{R}, P)$ . Note that this is the required approximation:

$$\begin{aligned} w(\mathcal{Q}' \cup \mathcal{R}'') &= w(\mathcal{Q}') + w(\mathcal{R}'') \leq (1 + \epsilon/3)w(\text{OPT}(\mathcal{R}', P')) + \epsilon w_{\text{appr}} \\ &\leq (1 + \epsilon/3)w(\mathcal{Q}) + \epsilon(1 + \epsilon/3)w(\mathcal{Q}) \leq (1 + \epsilon)w(\mathcal{Q}) \end{aligned}$$

Above we use the fact that  $\mathcal{Q}$  is also a set-cover for  $(\mathcal{R}', P')$ . Scaling by  $n/\epsilon w_{\text{approx}}$ , each set in  $\mathcal{R}'$  has weight at least 1, and weight of  $\text{OPT}(\mathcal{R}', P') = O(n/\epsilon)$ .  $\square$

Hence for the purpose of a  $(1 + \epsilon)$ -approximation, we can assume that the minimum weight of any region is 1 and the weight of the optimal set-cover is  $O(n/\epsilon)$ .

### 3 QPTAS for Weighted Pseudodisks in $\mathbb{R}^2$

Our main result in this section is:

**Theorem 3.1.** *Let  $\mathcal{R} = \{R_1, \dots, R_n\}$  be a set of  $n$  weighted  $\alpha$ -simple pseudodisks with minimum weight 1. Let  $P$  be a set of points in the plane, with no point lying on the boundary of any of the pseudodisks. Assume also that no pseudodisk in  $\text{OPT}(\mathcal{R}, P)$  has weight more than  $w(\text{OPT}(\mathcal{R}, P))/3$ . Then for any  $\delta > 0$ , there exists a curve  $\mathcal{C}$  such that*

- $w(\text{OPT}(\mathcal{R}, P_{\text{in}}(\mathcal{C}))) \leq (\frac{2}{3} + 3\delta)w(\text{OPT}(\mathcal{R}, P))$
- $w(\text{OPT}(\mathcal{R}, P_{\text{ext}}(\mathcal{C}))) \leq (\frac{2}{3} + 3\delta)w(\text{OPT}(\mathcal{R}, P))$
- $w(\text{OPT}(\mathcal{R}, P_{\text{in}}(\mathcal{C}))) + w(\text{OPT}(\mathcal{R}, P_{\text{ext}}(\mathcal{C}))) \leq (1 + 2\delta)w(\text{OPT}(\mathcal{R}, P))$
- *The complexity of  $\mathcal{C}$  is  $O(\frac{\alpha}{\delta^2} \log w(\text{OPT}(\mathcal{R}, P)))$ .*

A QPTAS for weighted pseudodisks follows from this theorem in similar manner to that of Adamaszek-Wiese [1, 2]. We first use Claim 2.4 to reduce the given instance of the set cover problem to an instance  $(\mathcal{R}, P)$  where the minimum weight of the regions is 1 and the weight of the optimal solution  $w(\text{OPT}(\mathcal{R}, P))$  is  $O(n/\epsilon)$ . Assume also that no pseudodisk in  $\text{OPT}(\mathcal{R}, P)$  has weight more than  $w(\text{OPT}(\mathcal{R}, P))/3$ . The input instance can be easily perturbed so that no point lies on the boundary of any region. Now by applying Theorem 3.1 with a given  $\delta$ , there exists a curve  $\mathcal{C}$  of complexity  $O(1/\delta^2 \log(n/\epsilon))$ . Thus by enumeration, there are  $O(n^{(1/\delta^2 \cdot \log n/\epsilon)^{O(1)}})$  such possible curves  $\mathcal{C}'$  (the proof of Theorem 3.1 shows that the vertices of any such  $\mathcal{C}'$  come from a polynomial-sized subset that can be computed in polynomial time) each giving two sub-problems  $(\mathcal{R}, P_{\text{in}}(\mathcal{C}'))$  and  $(\mathcal{R}, P_{\text{ext}}(\mathcal{C}'))$ . Thus, as  $\epsilon$  is a constant, the problem is QPT-partitionable, which together with Lemma 2.1 gives the required QPTAS. Finally, note that there can be at most 2 pseudodisks in the optimal solution with weight more than  $w(\text{OPT}(\mathcal{R}, P))/3$ , and one can simply guess (by enumerating the at most  $O(n^2)$  possibilities) these pseudodisks, and recurse on the sub-problem where the weight of the optimal solution is reduced by a constant-factor.

Towards proving the above theorem, we define structural decompositions for pseudodisks in  $\mathbb{R}^2$  called *core decompositions*. We will also use this notion for halfspaces in  $\mathbb{R}^3$ . Informally, given a set of pseudodisks, our goal is to shrink them in such a way that their union remains (almost) unchanged but the number of vertices<sup>4</sup> in the arrangement decreases. Denote by  $B_\tau$  a closed ball of radius  $\tau$ . We denote the Minkowski sum by  $\oplus$ .

**Definition 3.2 (Core Decomposition.)** Given  $\mathcal{R} = \{R_1, \dots, R_n\}$  and a  $\beta > 0$ , a set of regions  $\tilde{\mathcal{R}} = \{\tilde{R}_1, \dots, \tilde{R}_n\}$  is called a  $\beta$ -core decomposition of  $\mathcal{R}$  (and each  $\tilde{R}_i$  a *core* of  $R_i$ ) if (1)  $\tilde{R}_i \subseteq R_i$  for all  $i = 1, \dots, n$ , (2)  $\bigcup_i \tilde{R}_i \supseteq \bigcup_i R_i \setminus \bigcup_i (\partial R_i \oplus B_\beta)$ , and (3) each  $\tilde{R}_i$  is simply connected.

<sup>4</sup>By vertices we mean the intersection points of the boundaries of the pseudodisks.

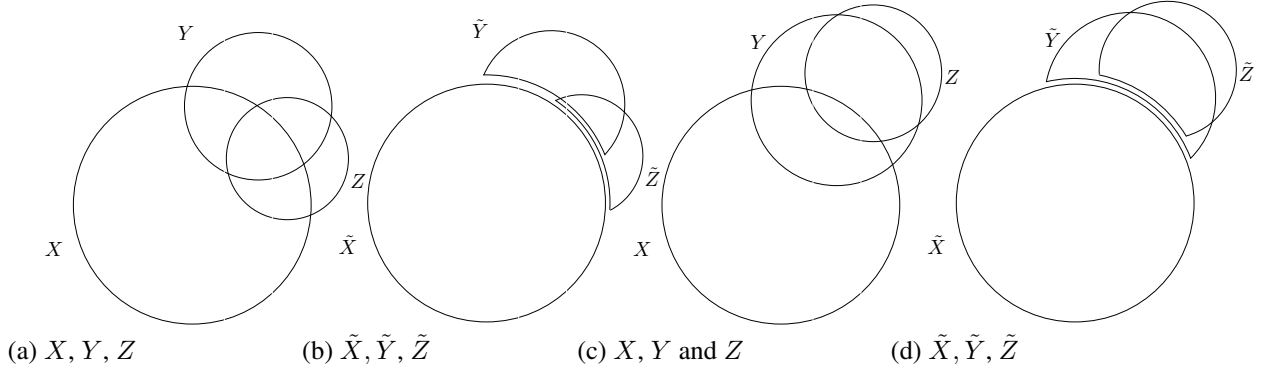


Figure 1: Pushing pseudodisks

Each disk  $\tilde{R}_i \in \tilde{\mathcal{R}}$  will be composed of pieces of boundaries of the disks in  $\mathcal{R}$ . The sequence of the endpoints of these pieces will be the vertices defining  $\tilde{R}_i$  (denote this sequence by  $v(\tilde{R}_i)$ , and its cardinality by  $|v(\tilde{R}_i)|$ ). The information needed to uniquely determine  $\tilde{R}_i$  then is the sequence of these pieces, or equivalently, the sequence of vertices defining  $\tilde{R}_i$ . In the following, we will use the term *core decomposition* to mean a  $\beta$ -core decomposition with a suitably small  $\beta > 0$  to be fixed later. The following two lemmas show the existence of core decompositions with specific properties.

**Lemma 3.3.** *Given a cover-free set  $\mathcal{R}$  of pseudodisks, a marked pseudodisk  $X \in \mathcal{R}$  (called the pusher) and a  $\beta > 0$ , there exists a  $\beta$ -decomposition  $\tilde{\mathcal{R}}$  of  $\mathcal{R}$  such that  $\tilde{X} = X$  and  $\tilde{R} \cap \tilde{X} = \emptyset$  for all  $R \neq X$  and  $\{\tilde{R} : R \in \mathcal{R}\}$  is a cover-free family of pseudodisks.*

*Proof.* Set  $\tilde{X} = X$ . For each  $R \in \mathcal{R} \setminus \{X\}$ , we compute a number  $gap(R) \in (0, \mu)$ , where  $\mu \in (0, \beta)$  is a suitably small number, and set  $\tilde{R} = \text{closure}(R \setminus (X \oplus B_{gap(R)}))$  (we say that  $X$  pushes  $R$  with gap  $gap(R)$ ). For any  $R \in \mathcal{R} \setminus \{X\}$ , let  $I_R$  be the interval  $R \cap \partial X$  on the boundary of  $X$ . As no pseudodisk in  $\mathcal{R}$  is completely contained in any other pseudodisk of  $\mathcal{R}$  (cover-free), the intervals  $I_R$  are well-defined. Consider the partial order  $\preceq$  on these intervals defined by inclusion ( $I_R \preceq I_S$  if  $I_R \subset I_S$ ). By a topological sorting of this partial order we can assign a distinct rank  $rank(R) \in \{1, \dots, n\}$  to each pseudodisk  $R \in \mathcal{R}$  such that if  $I_R \subset I_{R'}$  then  $rank(R) > rank(R')$ . We set  $gap(R) = \mu \frac{rank(R)}{n}$ .

Clearly each core is contained in its corresponding pseudodisk and for a small-enough  $\mu$ , it is simply connected. Also, since the points we may have removed from the union, due to the gaps, lie in  $X \oplus B_\mu$ , the second condition in the definition of a  $\beta$ -core decomposition is satisfied. The cores obtained are also cover-free because the union of cores cover the union of the original regions (except close to boundaries). Since the input set is cover-free each pseudodisk has a *free* portion that is not covered by others. The core corresponding to a pseudodisk then must cover the free portion in that pseudodisk which is not covered by the other cores. Thus no core is covered by the union of other cores.

It remains to show that the cores form pseudodisks. Let  $Y$  and  $Z$  be any two pseudodisks in  $\mathcal{R}$ , and we now finish the proof by showing that the boundaries of  $\tilde{Y}$  and  $\tilde{Z}$  intersect at most twice. The possible cases are the following: (1)  $I_Y \cap I_Z = \emptyset$ , (2)  $I_Y \preceq I_Z$  and (3)  $I_Y \cap I_Z \neq \emptyset$  and  $I_Y \not\preceq I_Z$ .

In case 1, since the intervals are disjoint,  $\partial \tilde{Y}$  and  $\partial \tilde{Z}$  do not have any new intersection that  $\partial Y$  and  $\partial Z$  did not have. They may have lost intersections lying in  $X$ . In any case,  $\partial \tilde{Y}$  and  $\partial \tilde{Z}$  intersect at most twice. In case 2,  $Y$  gets pushed with a smaller gap than  $Z$  and the situation is exactly as shown in Figures 1c and 1d. In case 3,  $Y$  and  $Z$  get pushed with different gaps and the situation is exactly as shown in Figures 1a and 1b.  $\square$

**Remark:** Note that for each pseudodisk  $R$  intersecting  $X$ , the boundary of  $\tilde{R}$  now has two new vertices corresponding to the two intersections of  $\partial R$  with  $\partial X$ . These vertices are slightly perturbed (and arbitrarily close) from the intersections because of the gap. We say that each such new vertex *corresponds* to the original intersection vertex between  $\partial R$  and  $\partial X$ . When the context is clear, we will not distinguish between this new vertex and the vertex it corresponds to.

Before we prove our next main result on core decompositions, we will need the following technical result. For clarity, vertex  $v$ , the intersection point of  $R_i$  and  $R_j$ , is written as  $(v, i, j)$ . Given  $\mathcal{R}$ , the *depth* of a vertex  $(v, i, j)$ , denoted  $d_v$ , is the total weight of the regions in  $\mathcal{R}$  containing  $(v, i, j)$  in the interior (thus it excludes the weight of  $R_i$  and  $R_j$ ).

**Claim 3.4.** *Let  $\mathcal{R} = \{R_1, \dots, R_n\}$  be a set of  $n$  weighted pseudodisks, and  $k > 0$  a given parameter. Assume  $R_i$  has weight  $w_i$ , and  $W = \sum_i w_i$ . Further let  $(v, i, j)$  denote a vertex  $v$  in the arrangement of  $\mathcal{R}$  defined by  $R_i$  and  $R_j$ . Then*

$$\sum_{\substack{(v,i,j) \text{ s.t.} \\ k \leq d_v < 2k}} \frac{w_i \cdot w_j}{w_i + w_j + k} = O(W)$$

*Proof.* The proof follows from melding the Clarkson-Shor technique with a charging argument. Let  $\mathcal{R}_1 \subset \mathcal{R}$  be the set of disks with weight  $w_i \geq 2k$ , and  $\mathcal{R}_2 = \mathcal{R} \setminus \mathcal{R}_1$ . Note that  $|\mathcal{R}_1| = O(W/k)$ , and that any vertex  $v$  with depth less than  $2k$  and defined by two disks in  $\mathcal{R}_1$  must lie on the boundary of the union of the regions in  $\mathcal{R}_1$ . This implies

$$\sum_{\substack{(v,i,j) \text{ s.t. } R_i, R_j \in \mathcal{R}_1 \\ k \leq d_v < 2k}} \frac{w_i \cdot w_j}{w_i + w_j + k} \leq \sum_{\substack{(v,i,j) \text{ s.t. } R_i, R_j \in \mathcal{R}_1 \\ k \leq d_v < 2k}} \min\{w_i, w_j\} = O\left(\sum_{R_i \in \mathcal{R}_1} w_i\right) = O(W) \quad (1)$$

where the second-last inequality follows from a charging argument: the number of vertices in the union for  $t$  pseudodisks is  $O(t)$  [30], and so each vertex can be assigned to one of its two disks such that each disk gets  $O(1)$  assigned vertices.

Set  $S = \mathcal{R}_1$ , and further add each disk  $R_i \in \mathcal{R}_2$  into  $S$  with probability  $p_i = w_i/4k$ . Then

$$\text{Expected union complexity of } S = E[O(|S|)] = O(|\mathcal{R}_1| + \sum_{R_i \in \mathcal{R}_2} \frac{w_i}{4k}) = O(W/k)$$

On the other hand, the expected number of vertices defined by the intersection of a disk in  $\mathcal{R}_2$  and a disk in  $\mathcal{R}_1$ , and of depth less than  $2k$ , that end up as vertices in the union of  $S$ :

$$\sum_{\substack{(v,i,j) \text{ s.t. } R_i \in \mathcal{R}_2, R_j \in \mathcal{R}_1 \\ k \leq d_v < 2k}} \frac{w_i}{4k} \prod_{\substack{R_l \text{ contains } v \\ R_l \in \mathcal{R}_2}} (1 - p_l) \geq \sum \frac{w_i}{4k} e^{-2 \sum_l w_l/4k} \geq \sum \frac{w_i}{4ek^2}$$

using the fact that  $1 - x \geq e^{-2x}$  for  $0 \leq x \leq 0.5$ , and that  $\sum w_l \leq 2k$  as all such  $R_l$  contain  $v$ , which has depth at most  $2k$ . The expected number of vertices in the union of  $S$  defined by two disks in  $\mathcal{R}_2$ , and of depth at most  $2k$ :

$$\sum_{\substack{(v,i,j) \text{ s.t. } R_i \in \mathcal{R}_2, R_j \in \mathcal{R}_2 \\ k \leq d_v < 2k}} \frac{w_i}{4k} \cdot \frac{w_j}{4k} \prod_{\substack{R_l \text{ contains } v \\ R_l \in \mathcal{R}_2}} (1 - p_l) \geq \sum \frac{w_i \cdot w_j}{16ek^2}$$



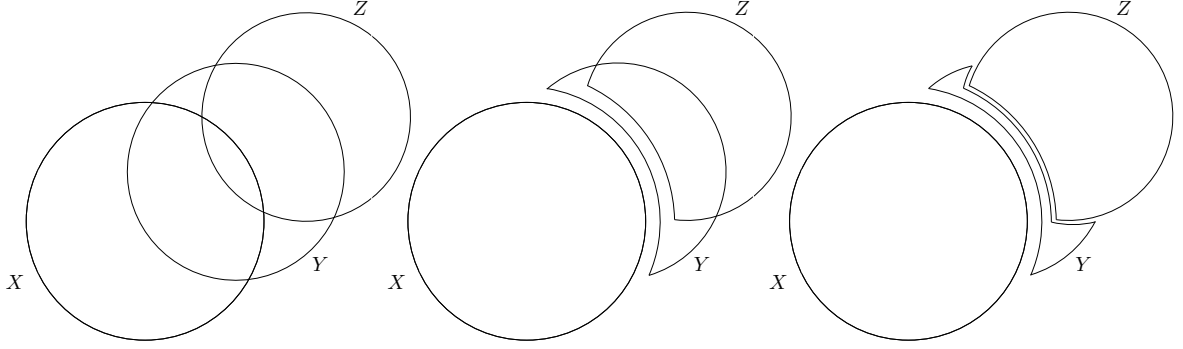


Figure 2:  $X$  is the first pusher,  $Z$  is the second pusher and  $Y$  is the third pusher.

Putting the lower- and upper-bounds together, we arrive at:

$$\sum_{(v,i,j) \text{ s.t. } R_i \in \mathcal{R}_2, R_j \in \mathcal{R}_1} \frac{w_i}{4ek} + \sum_{(v,i,j) \text{ s.t. } R_i \in \mathcal{R}_2, R_j \in \mathcal{R}_2} \frac{w_i \cdot w_j}{16ek^2} \leq O(W/k) \quad (2)$$

Finally,

$$\begin{aligned} \sum_{\substack{(v,i,j) \text{ s.t.} \\ k \leq d_v < 2k}} \frac{w_i \cdot w_j}{w_i + w_j + k} &= \sum_{\substack{(v,i,j) \\ R_i \in \mathcal{R}_2, R_j \in \mathcal{R}_1}} \frac{w_i \cdot w_j}{w_i + w_j + k} + \sum_{\substack{(v,i,j) \\ R_i \in \mathcal{R}_2, R_j \in \mathcal{R}_2}} \frac{w_i \cdot w_j}{w_i + w_j + k} \\ &\quad + \sum_{\substack{(v,i,j) \\ R_i \in \mathcal{R}_1, R_j \in \mathcal{R}_1}} \frac{w_i \cdot w_j}{w_i + w_j + k} \\ &\leq \sum_{\substack{(v,i,j) \\ R_i \in \mathcal{R}_2, R_j \in \mathcal{R}_1}} w_i + \sum_{\substack{(v,i,j) \\ R_i \in \mathcal{R}_2, R_j \in \mathcal{R}_2}} \frac{w_i \cdot w_j}{k} + O(W) = O(W) \end{aligned}$$

where the last inequality follows from Equation (2). □

**Lemma 3.5.** *Let  $\mathcal{R} = \{R_1, \dots, R_n\}$  be a set of  $n$  weighted pseudodisks, where  $R_i$  has weight  $w_i$ , and  $W = \sum_i w_i$ . Then there exists a core decomposition of  $\mathcal{R}$ , say the set  $\tilde{\mathcal{R}} = \{\tilde{R}_1, \dots, \tilde{R}_n\}$ , such that the pseudodisks in  $\tilde{\mathcal{R}}$  are pairwise disjoint, and*

$$\sum_i |v(\tilde{R}_i)| \cdot w_i = O(W \log W)$$

*Proof.* Recall that  $v(\tilde{R}_i)$  is the sequence of vertices defining  $\tilde{R}_i$  (which will be composed of pieces of boundaries of regions in  $\mathcal{R}$ ), and determines  $\tilde{R}_i$ .

The algorithm to construct  $\tilde{\mathcal{R}}$  is the following. Construct a permutation, say  $\pi$ , of  $\mathcal{R}$  randomly w.r.t. to the weight distribution of the disks as follows. Pick a random disk, where  $R_i$  is picked with probability  $w_i/W$ . Set this disk to be the first disk in the permutation, and recursively construct the rest of the permutation on the remaining disks. Let  $\pi_i$  denote the position of  $R_i$  in this permutation. Let  $\mathcal{R}_0 = \mathcal{R}$ . Apply Lemma 3.3 (with  $R_{\pi^{-1}(1)} \in \mathcal{R}_0$  as the pusher) to get a core decomposition, denoted by  $\mathcal{R}_1$ , of  $\mathcal{R}_0$ . Now apply Lemma 3.3

on  $\mathcal{R}_1$  (with  $\tilde{R}_{\pi^{-1}(2)} \in \mathcal{R}_1$  as the pusher) to get the set  $\mathcal{R}_2$ . Continuing iteratively by applying Lemma 3.3 with the successive core of each  $R_{\pi^{-1}(i)}$  in  $\mathcal{R}_i$  as the pusher, we get the set  $\tilde{\mathcal{R}} = \mathcal{R}_n$ . See Figure 2 for an example.

We have the following crucial fact:

**Claim 3.6.** *If  $v(\tilde{R}_l)$  has a vertex corresponding to  $(v, i, j)$ , then  $(v, i, j) \in R_l$  and furthermore,*

$$\begin{cases} \max\{\pi_i, \pi_j\} < \min_{R_m \ni v} \pi_m & \text{if } l \neq i, j \\ \pi_j < \pi_i & \text{if } l = i \\ \pi_i < \pi_j & \text{if } l = j \end{cases}$$

*Proof.* First consider the case when  $l \neq i, j$ . The proof relies on the observation that by the proof of Lemma 3.3, if  $v$  is a vertex in the arrangement of  $\mathcal{R}$ , and at any point a region containing  $v$  in its interior is used as a pusher, then any  $\tilde{R} \in \tilde{\mathcal{R}}$  cannot have a vertex that corresponds to  $v$ . Thus the only way a vertex corresponding to  $(v, i, j)$  can be part of the boundary of any  $\tilde{R}$  is if both the regions  $R_i$  and  $R_j$  occur earlier in  $\pi$  than any of the regions containing  $v$ . For the case when  $l = i$  (same for when  $l = j$ ), the proof follows from the fact that if at any point there are two regions  $X$  and  $Y$  and  $X$  is used as a pusher before  $Y$ , then the core of  $X$  cannot have a vertex that corresponds to a vertex defined by  $X$  and  $Y$  (see Figure 2).  $\square$

Then

$$\sum_i |v(\tilde{R}_i)| \cdot w_i = \sum_{(v,i,j)} \sum_{l \text{ s.t. } v \in v(\tilde{R}_l)} w_l = \sum_{(v,i,j)} \sum_{l \text{ s.t. } v \in R_l} X_{(v,i,j,l)} \cdot w_l$$

where the indicator variable  $X_{(v,i,j,l)} = 1$  iff  $v \in v(\tilde{R}_l)$  (more precisely, the vertex corresponding to  $v$  is in  $v(\tilde{R}_l)$ ). Using the above Claim and Claim 3.4, we calculate the expected value of the required bound:

$$\begin{aligned} \sum_i E[|v(\tilde{R}_i)|] \cdot w_i &= \sum_{(v,i,j)} \sum_{l \text{ s.t. } v \in R_l} E[X_{(v,i,j,l)}] \cdot w_l \\ &= \sum_{(v,i,j)} \sum_{l \text{ s.t. } v \in R_l} \Pr[v \in \tilde{R}_l] \cdot w_l \\ &= \sum_{(v,i,j)} \left( \Pr[v \in \tilde{R}_i] w_i + \Pr[v \in \tilde{R}_j] w_j + \sum_{\substack{i,j \neq l \\ v \in R_l}} \Pr[v \in \tilde{R}_l] w_l \right) \\ &= \sum_{(v,i,j)} \left( \frac{w_j}{w_i + w_j + d_v} \frac{w_i}{w_i + d_v} w_i + \frac{w_i}{w_i + w_j + d_v} \frac{w_j}{w_j + d_v} w_j \right. \\ &\quad \left. + \sum_{\substack{i,j \neq l \\ v \in R_l}} \left( \frac{w_j}{w_i + w_j + d_v} \frac{w_i}{w_i + d_v} + \frac{w_i}{w_i + w_j + d_v} \frac{w_j}{w_j + d_v} \right) w_l \right) \\ &= \sum_{(v,i,j)} \left( \left( \frac{w_j}{w_i + w_j + d_v} \frac{w_i}{w_i + d_v} \right) \cdot (w_i + \sum_{R_l \ni v} w_l) \right. \\ &\quad \left. + \left( \frac{w_i}{w_i + w_j + d_v} \frac{w_j}{w_j + d_v} \right) \cdot (w_j + \sum_{R_l \ni v} w_l) \right) \\ &= 2 \sum_{(v,i,j)} \frac{w_i \cdot w_j}{w_i + w_j + d_v} = 2 \sum_k \sum_{\substack{(v,i,j) \\ 2^k \leq d_v < 2^{k+1}}} \frac{w_i \cdot w_j}{w_i + w_j + d_v} = O(W \log W). \end{aligned}$$

where the last inequality follows from Claim 3.4.  $\square$

We can now finish the proof of Theorem 3.1.

*Proof of Theorem 3.1.* We use Lemma 3.5 to obtain a core  $\tilde{Q}$  for each  $Q \in \mathcal{Q} = \text{OPT}(\mathcal{R}, P)$  and we assign to  $\tilde{Q}$  the same weight as  $Q$ . Since the regions in  $\mathcal{Q}$  cover  $P$  their cores also cover  $P$ .<sup>5</sup> As before, we denote the number of vertices in  $\tilde{Q}$  by  $|v(\tilde{Q})|$ . By Lemma 3.5,  $\sum_{Q \in \mathcal{Q}} |v(\tilde{Q})|w(Q) = O(w(\mathcal{Q}) \log w(\mathcal{Q}))$ . We set  $\tau = C \cdot \frac{1}{\delta} \log w(\mathcal{Q})$  for some large enough constant  $C$ . Then by averaging,  $\sum_{Q \in \mathcal{Q}: |v(\tilde{Q})| > \tau} w(Q) < \delta w(\mathcal{Q})$ .

Let  $\mathcal{Q}_s = \{Q \in \mathcal{Q} : |v(\tilde{Q})| \leq \tau\}$  and let  $\tilde{\mathcal{Q}}_s = \{\tilde{Q} : Q \in \mathcal{Q}_s\}$ . The regions in  $\tilde{\mathcal{Q}}_s$  are  $\alpha\tau$ -simple since they have at most  $\tau$  sides and each of the sides is a portion of the boundary of a single  $\alpha$ -simple region in  $\mathcal{R}$ . These regions have a total weight of  $w(\mathcal{Q}_s)$ . Thus applying Theorem 2.2, we get separator  $\mathcal{C}$  so that the total weight of the regions of  $\mathcal{Q}_s$  whose cores lie in  $\text{interior}(\mathcal{C})$  ( $\text{exterior}(\mathcal{C})$ ) is at most  $\frac{2}{3}w(\mathcal{Q}_s)$ . Since the total weight of the regions in  $\mathcal{Q} \setminus \mathcal{Q}_s$  is at most  $\delta w(\mathcal{Q})$ , the total weight of all the cores that lie in  $\text{interior}(\mathcal{C})$  ( $\text{exterior}(\mathcal{C})$ ) is at most  $(\frac{2}{3} + \delta)w(\mathcal{Q})$ . Also, the total weight of the cores in  $\mathcal{Q}_s$  that intersect  $\mathcal{C}$  is at most  $\delta w(\mathcal{Q}_s)$ . Thus the total weight of all the cores in  $\mathcal{Q}$  that intersect  $\mathcal{C}$  is at most  $2\delta w(\mathcal{Q})$ . The complexity of  $\mathcal{C}$  is  $O(\alpha\tau/\delta) = O(\frac{\alpha}{\delta^2} \log w(\mathcal{Q}))$ , satisfying the fourth item in the statement of the theorem.

Let  $\mathcal{Q}_1$  ( $\mathcal{Q}_2$ ) be the set of regions whose cores are in  $\text{interior}(\mathcal{C})$  ( $\text{exterior}(\mathcal{C})$ ). Let  $\mathcal{Q}_3 = \mathcal{Q} \setminus \{\mathcal{Q}_1 \cup \mathcal{Q}_2\}$ . Observe that the cores of the regions in  $\mathcal{Q}_1 \cup \mathcal{Q}_3$  cover all the points in  $P_{in}(\mathcal{C})$  and therefore the regions in  $\mathcal{Q}_1 \cup \mathcal{Q}_3$  themselves cover the points in  $P_{in}(\mathcal{C})$ . Similarly the regions in  $\mathcal{Q}_2 \cup \mathcal{Q}_3$  cover the points in  $P_{ext}(\mathcal{C})$ . Therefore,

$$w(\text{OPT}(\mathcal{R}, P_{in})) \leq w(\mathcal{Q}_1 \cup \mathcal{Q}_3) = w(\mathcal{Q}_1) + w(\mathcal{Q}_3) \leq (\frac{2}{3} + 3\delta)w(\mathcal{Q})$$

This proves the first item in the statement of the theorem. The second item is proved analogously. For the third item, we combine the inequalities  $w(\text{OPT}(\mathcal{R}, P_{in})) \leq w(\mathcal{Q}_1) + w(\mathcal{Q}_3)$  and  $w(\text{OPT}(\mathcal{R}, P_{ext})) \leq w(\mathcal{Q}_2) + w(\mathcal{Q}_3)$ . We get

$$w(\text{OPT}(\mathcal{R}, P_{in})) + w(\text{OPT}(\mathcal{R}, P_{ext})) \leq w(\mathcal{Q}_1) + w(\mathcal{Q}_2) + 2w(\mathcal{Q}_3) \leq w(\mathcal{Q}) + w(\mathcal{Q}_3) \leq (1 + 2\delta)w(\mathcal{Q})$$

That proves the third item.  $\square$

**Remark:** The above QPTAS can be extended to work for more general regions called *non-piercing regions* or *r-admissible regions*. For this only Lemma 3.3 needs to be extended to work for these regions. In this case, a region  $R$  may intersect the boundary of pusher  $X$  in more than one interval. To ensure that after pushing the new regions are still non-piercing, different gaps are required in different intervals for the same region  $R$ . This makes it technically more complicated. The details are given in Appendix B.

## 4 QPTAS for Weighted Halfspaces in $\mathbb{R}^3$

Let  $\mathcal{H} = \{H_1, \dots, H_n\}$  be a set of halfspaces in  $\mathbb{R}^3$  where the halfspace  $H_i$  has weight  $w_i \geq 0$  and  $W$  total weight. Let  $P$  be a set of points in  $\mathbb{R}^3$ . Given  $\mathcal{H}$  and  $P$ , we show that the problem of computing a

<sup>5</sup>Since no point lies on the boundary of any of the regions, there is a suitable choice of  $\beta$  so that using  $\beta$ -core decompositions, we do not miss any of the points.

subset of  $\mathcal{H}$  of minimum weight whose union covers  $P$  is QPT-partitionable, and then Lemma 2.1 implies the QPTAS.

Consider the optimal solution  $\text{OPT} = \text{OPT}(\mathcal{H}, P)$  for the problem, and let  $W$  be the total weight of the halfspaces in  $\text{OPT}$ . For any halfspace  $H$ , define  $\bar{H}$  to be other halfspace defined by its boundary  $\partial H$  i.e.,  $\bar{H} = \text{closure}(\mathbb{R}^3 \setminus H)$ . For any set of halfspaces  $S$ , define  $\bar{S} = \{\bar{H} : H \in S\}$ .

**Lemma 4.1.** *If  $\bigcup_{H \in \text{OPT}} H = \mathbb{R}^3$ , then one can compute  $\text{OPT}(\mathcal{H}, P)$  in polynomial time.*

*Proof.* If  $\bigcup_{H \in \text{OPT}} H = \mathbb{R}^3$ , then by definition  $\bigcap_{H \in \text{OPT}} \bar{H} = \emptyset$ . By Helly's theorem [24] applied to the set of convex regions in  $\bar{\text{OPT}}$ , it follows that then there must be a subset  $\text{OPT}' \subset \text{OPT}$  of at most 4 halfspaces such that  $\bigcap_{H \in \text{OPT}'} \bar{H} = \emptyset$ . In other words,  $\text{OPT}'$  covers  $\mathbb{R}^3$ . As  $\text{OPT}$  was a minimal-weight set cover, it follows that  $|\text{OPT}| \leq 4$ . By enumerating all 4-tuples of halfspaces in  $\mathcal{H}$ , one can compute the optimal set-cover in polynomial time.  $\square$

From now on we assume that there is a point  $o$  that does not lie in any of the halfspaces in  $\text{OPT}$  (say the origin). We will also assume without loss of generality that the intersection of halfspaces in  $\bar{\text{OPT}}$  is a bounded polytope. This can be easily done by adding to the input four halfspaces with weight 0 which do not contain any of the points in  $P$  whose complements intersect in a bounded simplex. These four dummy halfspaces can then be included in any optimal solution without affecting the weight of the solution. Note also that each halfspace  $H \in \text{OPT}$  must be part of some facet (in fact, a unique facet) of this polytope; otherwise  $H$  is contained in the union of  $\text{OPT} \setminus \{H\}$ , contradicting the set-cover minimality of  $\text{OPT}$ .

We now define a core decomposition for the halfspaces in  $\text{OPT}$  that allows a cheap balanced polyhedral separator.

Consider the set system in which the base set are the halfspaces in  $\text{OPT}$  and subsets are defined by taking any segment  $ox$  with one end-point at  $o$  and taking the set of halfspaces whose boundaries intersect the segment. More formally, for any  $x \in \mathbb{R}^3$ , let  $R_x = \{H \in \text{OPT} : \partial H \cap ox \neq \emptyset\}$ . We now define  $\mathcal{R}$  as the set  $\{R_x : x \in \mathbb{R}^3\}$ . Consider the weighted set system  $(\mathcal{H}, \mathcal{R})$ , where the weight of  $w(R)$  of any  $R \in \mathcal{R}$  is the sum of the weights of the halfspaces in  $R$ .

**Lemma 4.2.** *The VC-dimension of  $(\mathcal{H}, \mathcal{R})$  is at most 3.*

*Proof.* For two distinct points  $x$  and  $y$  that lie in the same cell of the arrangement of the halfspaces in  $\text{OPT}$ ,  $R_x = R_y$ . So for any subset of  $\text{OPT}$  of size  $k$ , the number of induced subsets is at most the number of cells in an arrangement of these  $k$  halfspaces, which is at most  $\binom{k}{3} + \binom{k}{2} + \binom{k}{1} + \binom{k}{0}$ . For  $k = 4$ , this number is less than  $2^k$ , implying that no subset of size 4 is shattered. Thus the VC-dimension this set system is at most 3.  $\square$

Thus, by the  $\epsilon$ -net theorem [21], there is an  $\epsilon$ -net for this set system of size  $O(\frac{1}{\epsilon} \log \frac{1}{\epsilon})$ . Let  $N$  be an  $\epsilon$ -net for this set system for a value of  $\epsilon$  to be fixed later. As before, we will assume that the intersection of halfspaces in  $\bar{N}$  is a bounded polytope  $\mathcal{P}$ . This can be ensured by including in  $N$  the dummy halfspaces.

For any set  $S \subset \mathbb{R}^3$ , define  $\text{cone}(S)$  to be the set  $\{\lambda x : x \in S, \lambda \geq 0\}$ . For any halfspace  $H \in N$ , we define the core of  $H$  to be  $\tilde{H} = \text{cone}(f) \cap H$  where  $f$  is the facet of  $\mathcal{P}$  corresponding to  $H$  i.e., the facet contained in  $\partial H$ . Note that each halfspace in  $N$  (with the exception of dummy halfspaces) has a unique facet of  $\mathcal{P}$  corresponding to it. For any halfspace  $H \in \text{OPT} \setminus N$ , we defined the core as  $\tilde{H} = \mathcal{P} \cap H$ . The core of each halfspace is clearly contained in the halfspace and the union of these cores is clearly the same as the union of the halfspaces in  $\text{OPT}$ .

We now assign a weight to each of the facets of  $\mathcal{P}$  by distributing the weights of the halfspaces in  $\text{OPT}$  to the facets so that the total weight of the faces is the same as the total weight of the halfspaces. The weight of each halfspace in  $N$  is assigned to facet corresponding it. For a halfspace in  $H \in \text{OPT} \setminus N$ , we distribute its weight equally among all faces  $f$  s.t.  $\text{cone}(f)$  intersects the core  $\tilde{H}$  of  $H$ .

The 1-skeleton of  $\mathcal{P}$  is a planar graph  $G$  and we have assigned weights to its faces. Let  $n'$  denote the number of vertices in this graph; note that  $n' = O(|N|) = O(\frac{1}{\epsilon} \log \frac{1}{\epsilon})$ . By [26], there exists in this graph a cycle separator  $\mathcal{C}$  of size  $O(\sqrt{n'})$  so that the total weight of the faces in the interior (exterior) of  $\mathcal{C}$  is at most two thirds of the total weight. We show that the polytope  $\hat{\mathcal{C}} = \text{cone}(\mathcal{C})$  is the desired cheap balanced separator for the cores we have defined.  $\hat{\mathcal{C}}$  splits  $\mathbb{R}^3$  into two connected pieces whose closures we call the interior and the exterior of  $\hat{\mathcal{C}}$ . The choice is arbitrary. Note that  $\text{interior}(\hat{\mathcal{C}}) \cap \text{exterior}(\hat{\mathcal{C}}) = \hat{\mathcal{C}}$ .

First note that for each core that lies in the interior (exterior), the weight of the corresponding halfspaces is distributed only among the faces of  $\mathcal{P}$  lying in the interior (exterior) of  $\hat{\mathcal{C}}$ . Hence the total weight of all cores that lie in the interior (exterior) of  $\hat{\mathcal{C}}$  is at most two thirds of the total weight of all halfspaces.

We now need to bound the total weight of the cores that cross  $\hat{\mathcal{C}}$ . None of the cores of the halfspaces in  $N$  cross  $\hat{\mathcal{C}}$ . Consider a halfspace  $H \in \text{OPT} \setminus N$ . Its core is defined as  $H \cap \mathcal{P}$ . If this core intersects  $\hat{\mathcal{C}}$ , then  $H \cap \partial\mathcal{P}$  intersects  $\mathcal{C}$ . It follows that  $H$  intersects an edge of  $\mathcal{C}$  and thus must contain a vertex  $v$  of  $\mathcal{C}$ . In other words,  $\partial H$  intersects the segment  $ov$ . However, since  $N$  is an  $\epsilon$ -net, for any vertex  $v$ , the total weight of halfspaces whose boundaries intersects  $ov$  is at most  $\epsilon W$ . Since  $\mathcal{C}$  has  $O(\sqrt{n'})$  vertices, the total weight of all cores crossing  $\hat{\mathcal{C}}$  is  $O(\sqrt{n'} \cdot \epsilon W)$ . We set  $\epsilon = A\delta^2 / \log \delta^{-2}$  for a suitable constant  $A$  so that  $\mathcal{C}$  has  $O(\frac{1}{\delta} \log \frac{1}{\delta})$  vertices and the total weight of cores intersecting is at most  $\delta W$ .

Finally, observe that the complexity of  $\hat{\mathcal{C}}$  is determined by the complexity of  $\mathcal{C}$ , and the point  $o$ . The vertices of  $\mathcal{C}$  are determined by intersections of 3 halfspaces of  $\mathcal{H}$ , and so there are  $O(n^3)$  choices for each vertex of  $\mathcal{C}$ . To guess the point  $o$ , it suffices to guess the cell of the arrangement of  $\mathcal{H}$  in which it lies (there are  $O(n^3)$  such choices), and pick any point in that cell.

**Remark:** It may appear that the set cover problem for halfspaces may be reduced to the problem for pseudodisks using techniques used in [25]. Unfortunately, that does not work because (i) we are in the weighted setting and (ii) because we cannot tolerate the loss of a constant factor when looking for a  $(1 + \epsilon)$ -approximation algorithm. It is also tempting to think that the technique used for halfspaces may be used for pseudodisks in the plane. That would mean taking an  $\epsilon$ -net  $N$  for a suitable range space and then defining the core for each pseudodisks  $R \notin N$  by removing from  $R$  the portion of it covered by the union of pseudodisks in  $N$ . However, the problem in doing this is that the resulting cores may not be connected. This causes problems because if the cores are not connected then the cores not intersecting a separator curve  $\mathcal{C}$  may still cover points in both  $\text{interior}(\mathcal{C})$  and  $\text{exterior}(\mathcal{C})$ .

## 5 Conclusion

In this paper we demonstrated the versatility of separator-based algorithmic design on a problem seemingly unrelated to the packing problems for which the separator had previously been successfully applied. Getting a polynomial-time approximation scheme for the set-cover problem for weighted pseudodisks in the plane and weighted halfspaces in  $\mathbb{R}^3$  remains a very interesting open problem.

## References

- [1] A. Adamaszek and A. Wiese. Approximation schemes for maximum weight independent set of rectangles. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2013)*, 2013.
- [2] A. Adamaszek and A. Wiese. A QPTAS for maximum weight independent set of polygons with polylogarithmically many vertices. In *SODA*, 2014.
- [3] Pankaj K. Agarwal, Esther Ezra, and Micha Sharir. Near-linear approximation algorithms for geometric hitting sets. *Algorithmica*, 63(1-2):1–25, 2012.
- [4] Pankaj K. Agarwal and Micha Sharir. Davenport-schinz sequences and their geometric applications, 1998.
- [5] Boris Aronov, Esther Ezra, and Micha Sharir. Small-size  $\epsilon$ -nets for axis-parallel rectangles and boxes. *SIAM J. Comput.*, 39(7):3248–3282, 2010.
- [6] Nikhil Bansal and Kirk Pruhs. The geometry of scheduling. In *FOCS*, pages 407–414, 2010.
- [7] P. Carmi, M. Katz, and N. Lev-Tov. Covering points by unit disks of fixed location. In *ISAAC*, pages 644–655, 2007.
- [8] Timothy M. Chan and Elyot Grant. Exact algorithms and apx-hardness results for geometric packing and covering problems. *Comput. Geom.*, 47(2):112–124, 2014.
- [9] Timothy M. Chan, Elyot Grant, Jochen Könnemann, and Malcolm Sharpe. Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In *SODA*, 2012.
- [10] Bernard Chazelle and Joel Friedman. A deterministic view of random sampling and its use in geometry. *Combinatorica*, 10(3):229–249, 1990.
- [11] Victor Chepoi and Stefan Felsner. Approximating hitting sets of axis-parallel rectangles intersecting a monotone curve. *Comput. Geom.*, 46(9):1036–1041, 2013.
- [12] K. Clarkson and K. Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete Comput. Geom.*, 37:43–58, 2007.
- [13] Kenneth L. Clarkson and Kasturi R. Varadarajan. Improved approximation algorithms for geometric set cover. In *Symposium on Computational Geometry*, pages 135–141, 2005.
- [14] Francisco Claude, Reza Dorrigiv, Stephane Durocher, Robert Fraser, Alejandro López-Ortiz, and Alejandro Salinger. Practical discrete unit disk cover using an exact line-separable algorithm. In *ISAAC*, pages 45–54, 2009.
- [15] Gruia Călinescu, Ion I. Mandoiu, Peng-Jun Wan, and Alexander Z. Zelikovsky. Selecting forwarding neighbors in wireless ad hoc networks. *Mob. Netw. Appl.*, 9(2):101–111, 2004.
- [16] Gautam K. Das, Robert Fraser, Alejandro López-Ortiz, and Bradford G. Nickerson. On the discrete unit disk cover problem. In *WALCOM*, pages 146–157, 2011.
- [17] Thomas Erlebach and Erik Jan van Leeuwen. Approximating geometric coverage problems. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms, SODA '08*, 2008.
- [18] Jacob Fox and János Pach. Computing the independence number of intersection graphs. In *SODA*, pages 1161–1165, 2011.
- [19] Sarel Har-Peled. Being fat and friendly is not enough. *CoRR*, abs/0908.2369, 2009.

- [20] Sariel Har-Peled. Quasi-polynomial time approximation scheme for sparse subsets of polygons. In *Symposium on Computational Geometry*, to appear, 2014.
- [21] D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Comput. Geom.*, 2:127–151, 1987.
- [22] Dorit S. Hochbaum and Wolfgang Maass. Approximation schemes for covering and packing problems in robotics and vlsi. In *STACS*, pages 55–62, 1984.
- [23] J. Matousek. Efficient partition trees. *Discrete & Computational Geometry*, 8:315–334, 1992.
- [24] J. Matousek. *Lectures in Discrete Geometry*. Springer-Verlag, New York, NY, 2002.
- [25] J. Matousek, R. Seidel, and E. Welzl. How to net a lot with little: Small epsilon-nets for disks and halfspaces. In *Proceedings of Symposium on Computational Geometry*, pages 16–22, 1990.
- [26] Gary L. Miller. Finding small simple cycle separators for 2-connected planar graphs. *J. Comput. Syst. Sci.*, 32(3):265–279, 1986.
- [27] Nabil H. Mustafa. *Approximation of Points: Combinatorics and Algorithms*. Habilitation thesis, University of Paris-Est, 2013.
- [28] Nabil H. Mustafa and Saurabh Ray. Improved results on geometric hitting set problems. *Discrete & Computational Geometry*, 44(4):883–895, 2010.
- [29] S. Narayanappa and P. Vojtechovský. An improved approximation factor for the unit disk covering problem. In *CCCG*, 2006.
- [30] Janos Pach and Micha Sharir. *Combinatorial Geometry And Its Algorithmic Applications: The Alcala Lectures*. Mathematical Surveys And Monographs. American Mathematical Society, 2009.
- [31] R. Raz and M. Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proceedings of STOC*, pages 475–484, 1997.
- [32] Luca Trevisan. Non-approximability results for optimization problems on bounded degree instances. In *Proceedings of the Thirty-third Annual ACM Symposium on Theory of Computing*, STOC '01, pages 453–461, 2001.
- [33] K. Varadarajan. Weighted geometric set cover via quasi uniform sampling. In *Proceedings of ACM Symposium on Theory of Computing*, 2010.

## A OPTAS for Uniform Pseudodisks

For the uniform pseudodisk case, there is an easier proof, which we present now.

**Lemma A.1.** *Given a set of pseudodisks  $\mathcal{R} = \{R_1, \dots, R_n\}$  and a parameter  $\eta > 0$ , there exists a core decomposition of  $\mathcal{R}$ , say the set  $\tilde{\mathcal{R}} = \{\tilde{R}_1, \dots, \tilde{R}_n\}$ , such that*

1. *the number of intersections in the arrangement induced by  $\tilde{\mathcal{R}}$  is  $O(1/\eta^2 + \eta n^2)$ , and*
2. *the number of vertices of each pseudodisk in  $\tilde{\mathcal{R}}$  is  $O(1/\eta^2)$ .*

*Proof.* Construct an  $\eta$ -net, say the set  $Q = \{Q_0, \dots, Q_t\}$ , for  $\mathcal{R}$ . By the result of Clarkson-Varadarajan [13],  $Q$  has size  $O(1/\eta)$ .

Let  $\mathcal{R}_0 = \mathcal{R}$ . Apply Lemma 3.3 (with  $Q_0 \in \mathcal{R}_0$  as the pusher) to get a core decomposition, denoted by  $\mathcal{R}_1$ , of  $\mathcal{R}_0$ . Now apply Lemma 3.3 on  $\mathcal{R}_1$  (with  $\tilde{Q}_1 \in \mathcal{R}_1$  as the pusher) to get the set  $\mathcal{R}_2$ . Continuing iteratively by applying Lemma 3.3 with the successive core of each  $Q_i$  in  $\mathcal{R}_i$  as the pusher, we get the set  $\mathcal{R}_t$ . Replace all the cores of pseudodisks  $R \in Q$  in  $\mathcal{R}_t$  by  $R$  to get the set  $\tilde{\mathcal{R}}$ .

Observe that each  $\tilde{R} \in \tilde{\mathcal{R}}$  is disjoint from each object in  $Q$  (it became disjoint from  $Q_i$  latest at the  $i$ -th iteration). As  $Q$  was an  $\eta$ -net, any point in the plane not covered by the union of  $Q$  has depth at most  $\eta n$ . Recall that by the Clarkson-Shor technique, as pseudodisks have linear union complexity, the maximum number of vertices at depth at most  $k$  is  $O(nk)$ . Therefore the total number of intersections in  $\tilde{\mathcal{R}}$  is

$$O(|Q|^2 + |\mathcal{R}_t| \cdot \eta n) = O(1/\eta^2 + n \cdot \eta n)$$

This proves condition 1.

For condition 2., from the proof of Lemma 3.3, each boundary vertex of any core object corresponds to a vertex of the arrangement induced by the objects in  $Q$ . As every pair of pseudodisks can intersect at most twice, there are  $O(1/\eta^2)$  vertices in the arrangement of  $Q$ .  $\square$

We now give a proof of the existence of the separator for uniformly weighted pseudodisks.

**Lemma A.2.** *Given a set of  $\mathcal{R}$  of  $n$  uniformly weighted  $\alpha$ -simple pseudodisks (each with weight 1) and set  $P$  of points in the plane, no point lying on the boundary of any of the regions, and any parameter  $\delta$ , there exists a curve  $\mathcal{C}$  such that*

- $w(\text{OPT}(\mathcal{R}, P_{in}(\mathcal{C}))) \leq (\frac{2}{3} + \delta)w(\text{OPT}(\mathcal{R}, P))$
- $w(\text{OPT}(\mathcal{R}, P_{ext}(\mathcal{C}))) \leq (\frac{2}{3} + \delta)w(\text{OPT}(\mathcal{R}, P))$
- $w(\text{OPT}(\mathcal{R}, P_{in}(\mathcal{C}))) + w(\text{OPT}(\mathcal{R}, P_{ext}(\mathcal{C}))) \leq (1 + \delta)w(\text{OPT}(\mathcal{R}, P))$
- *the complexity of  $\mathcal{C}$  is  $O(\log^5 n / \delta^5)$*

where  $w(\mathcal{S})$  denotes the total weight of the regions in  $\mathcal{S}$ .

*Proof.* Consider the set  $\mathcal{Q} = \text{OPT}(\mathcal{R}, P)$ . We apply Lemma A.1 to  $\mathcal{Q}$  and get a core decomposition with a core  $\tilde{Q}$  for each  $Q \in \mathcal{Q}$ . Since, the regions in  $\mathcal{Q}$  cover  $P$ , their cores also cover  $P$ .<sup>6</sup> By the property of the

<sup>6</sup>Since no point lies on the boundary of any of the regions, there is suitable choice of  $\beta$  so that using  $\beta$ -core decompositions, we do not miss any of the points.



core decomposition of Lemma A.1, we have the total number of intersections to be  $m = O(\delta^2 n^2 / \log^2 n)$ , and each core has  $O(\log^4 n / \delta^4)$  vertices. Since each of the curves forming the boundary of any core is composed of at most  $\alpha$   $x$ -monotone curves, the cores are  $\alpha_1$ -simple for  $\alpha_1 = O(\alpha \log^4 n / \delta^4)$ . Applying Theorem 2.3 to these cores with parameter  $r = \alpha^2 \log^{10} n / \delta^{10}$ , we get a curve  $\mathcal{C}$ , where

$$\begin{aligned} \text{Number of cores intersected by } \mathcal{C} &= O(\sqrt{m + \alpha_1^2 n^2 / r}) = O(\sqrt{\frac{\delta^2 n^2}{\log^2 n}}) = O(\frac{\delta n}{\log n}) \\ \text{Complexity of } \mathcal{C} &= O(\sqrt{r + \frac{mr^2}{\alpha_1^2 n^2}}) = O(\sqrt{\frac{\log^{10} n}{\delta^{10}}}) = O(\frac{\alpha \log^5 n}{\delta^5}) \end{aligned}$$

Let  $\mathcal{Q}_1$  ( $\mathcal{Q}_2$ ) be the set of regions whose cores are in  $\text{interior}(\mathcal{C})$  ( $\text{exterior}(\mathcal{C})$ ). Let  $\mathcal{Q}_3 = \mathcal{Q} \setminus \{\mathcal{Q}_1 \cup \mathcal{Q}_2\}$ . Observe that the cores of the regions in  $\mathcal{Q}_1 \cup \mathcal{Q}_3$  cover all the points in  $P_{in}$  and therefore the regions in  $\mathcal{Q}_1 \cup \mathcal{Q}_3$  themselves cover the points in  $P_{in}$ . Similarly the regions in  $\mathcal{Q}_2 \cup \mathcal{Q}_3$  cover the points in  $P_{ext}$ . Theorem 2.2 guarantees that  $w(\mathcal{Q}_1)$  and  $w(\mathcal{Q}_2)$  are at most  $\frac{2}{3}w(\mathcal{Q})$  and  $w(\mathcal{Q}_3) \leq \delta w(\mathcal{Q})$ . Therefore,

$$w(\text{OPT}(\mathcal{R}, P_{in})) \leq w(\mathcal{Q}_1 \cup \mathcal{Q}_3) = w(\mathcal{Q}_1) + w(\mathcal{Q}_3) \leq (\frac{2}{3} + \delta)w(\mathcal{Q})$$

This proves the first item in the statement of the theorem. The second item is proved analogously. For the third item, we combine the inequalities  $w(\text{OPT}(\mathcal{R}, P_{in})) \leq w(\mathcal{Q}_1) + w(\mathcal{Q}_3)$  and  $w(\text{OPT}(\mathcal{R}, P_{ext})) \leq w(\mathcal{Q}_2) + w(\mathcal{Q}_3)$ . We get

$$w(\text{OPT}(\mathcal{R}, P_{in})) + w(\text{OPT}(\mathcal{R}, P_{ext})) \leq w(\mathcal{Q}_1) + w(\mathcal{Q}_2) + 2w(\mathcal{Q}_3) \leq w(\mathcal{Q}) + w(\mathcal{Q}_3) \leq (1 + \delta)w(\mathcal{Q})$$

That proves the third item. Lemma A.1 also gives a set  $\Gamma(\tilde{Q})$  for each  $Q \in \mathcal{Q}$ . Let  $Y = \cup_{Q \in \mathcal{Q}} \Gamma(\tilde{Q})$ . The separator  $\mathcal{C}$  given by Theorem 2.3 can be described by a sequence of  $O(\alpha \log^5 n / \delta^5)$  curves in  $Y$  and  $O(\alpha \log^5 n / \delta^5)$  additional bits. Since, by Lemma A.1 each curve in  $Y$  can be described by a sequence of at most three curves in  $\Gamma(\mathcal{R})$  and a constant number of bits,  $\mathcal{C}$  can also be described by a sequence of  $O(\alpha \log^5 n / \delta^5)$  curves in  $\Gamma(\mathcal{R})$  and  $O(\alpha \log^5 n / \delta^5)$  additional bits of information.  $\square$

Now the existence of  $\mathcal{C}$  immediately implies that the problem is QPT-partitionable, which together with Lemma 2.1 yields a QPTAS for the case with uniform weights.

## B OPTAS for $r$ -admissible regions

Two regions  $A$  and  $B$  are said to be non-piercing if they are simply connected, their boundaries intersect at most a finite number of times, and the regions  $A \setminus B$  and  $B \setminus A$  are connected. A finite set of regions is said to be non-piercing if they are pairwise non-piercing. Figure 8 shows three regions that form a non-piercing set. The intersection of any two regions  $A$  and  $B$  that are non-piercing consists of a disjoint union of *lenses* formed by the regions. Each lens is a connected component of  $A \cap B$  which lies between two intersection points of  $\partial A$  and  $\partial B$  that are consecutive along both the boundaries. The shaded areas in Figure 3 show the lenses in the intersection of a non-piercing pair of regions. Figure 5 shows a modification of the region  $B$  in Figure 3 so that one of the lenses is bypassed. The boundary of  $B$  is modified so that the portion of  $\partial B$  that lies inside  $A$  is replaced by the dashed curve shown in the figure that is arbitrarily close to  $\partial A$  but outside it. We will use such operations of *bypassing a lens* in the proof of the next lemma.

**Lemma B.1.** *Given a set  $\mathcal{R}$  of non-piercing regions, and a marked region  $X \in \mathcal{R}$  (called the pusher), there exists a core decomposition  $\tilde{\mathcal{R}}$  of  $\mathcal{R}$  such that i)  $\tilde{X} = X$  and  $\tilde{R} \cap \tilde{X} = \emptyset$  for all  $R \neq X$  and ii)  $\tilde{\mathcal{R}}$  forms a family of non-piercing regions.*

*Proof.* The idea for the proof is the same as with pseudodisks. We set  $\tilde{X} = X$  and  $X' = X \oplus B_\mu$  for a suitably small  $\mu$ . For sufficiently small  $\mu$ ,  $X'$  forms a non-piercing family with the regions in  $\mathcal{R} \setminus \{X\}$ .

The intersection of each region in  $\mathcal{R} \setminus \{X\}$  forms a set of intervals on the boundary of  $X$ . Let  $\mathcal{I}$  be the set of all these intervals corresponding to all the regions. We then consider the partial order among them defined by inclusion, just as we did in the case of pseudodisks. This time, instead of assigning ranks to the pseudodisks, we assign distinct ranks in the range 1 to  $m$  to the intervals, where  $m$  is the total number of intervals in  $\mathcal{I}$ . Different intervals corresponding to the same region are given different ranks by this procedure.

Unlike with pseudodisks, we do not push a region  $R$  with the single gap. Instead we push it with different gaps along the different intervals in which it intersects  $\partial X$ . The gap along an interval  $I$  is  $\text{gap}(I) = \mu \frac{\text{rank}(I)}{m}$ . To better understand the pushing procedure, we imagine pushing along the intervals in  $\mathcal{I}$  one by one in decreasing order of their ranks.

Figures 6 and 7 show the effect of pushing the region  $A$  along the interval  $I = [a, b]$  which is one of two intervals on  $\partial X$  in which  $A$  intersects  $\partial X$ . The part of  $\partial A$  inside  $X$  joining  $a$  and  $b$  is replaced by the dashed curve shown in Figure 7 that lies on the boundary of  $X \oplus B_{\text{gap}(I)}$ .

We argue that this change to a region  $A$  keeps the set of regions non-piercing. That is, the modified  $A$  forms a non-piercing pair with every other object  $B$  (which remains unchanged). To see this, imagine the change in  $A$  as the net result of a sequence of changes. Within  $X$ ,  $A$  may form 0 or more lenses with  $B$ . We first make  $A$  bypass each of these lenses one by one. As discussed before, each of these changes keeps  $A$  and  $B$  non-piercing. After this, the number of intersections between the boundaries of  $A$  and  $B$  within  $X$  is at most 1. At this point we may move the boundary of  $A$  to its final position on the boundary of  $X \oplus B_{\text{gap}(I)}$ . Observe that this final step does not change the number of intersections between  $\partial A$  and  $\partial B$  unless  $B$  intersects  $\partial X$  in a sub-interval  $I' \subset I$ , as shown in Figure 4. However, in that case, we push  $B$  along the interval  $I'$  with a larger gap, before pushing  $A$  along  $I$  since  $I'$  gets a larger rank than  $I$ . Thus, the modified  $A$  bypasses some of the lenses with other objects (one with  $B$ , two with  $C$  and one with  $D$  in Figure 6) but this does not affect the non-piercing property of the family of regions. As another example, Figures 8 and 9 show the effect of pushing two regions  $A$  and  $B$  with a third one  $X$ .

□

**Remark:** Note that, as in the case of pseudodisks, for each  $R \in \mathcal{R}$ , the boundary of  $\tilde{R}$  gains as many

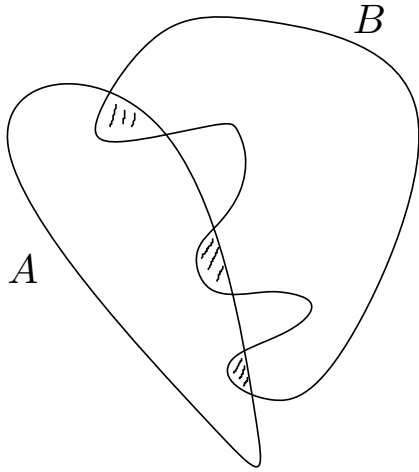


Figure 3: Lenses in the intersection of two non-piercing regions.

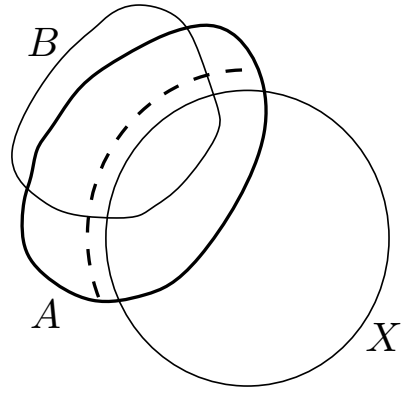


Figure 4: The algorithm pushes  $B$  before  $A$ . Therefore the situation shown in the figure cannot happen.

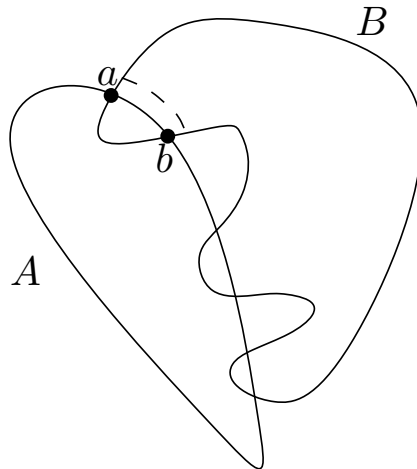


Figure 5:  $B$  is modified to bypass a lens.

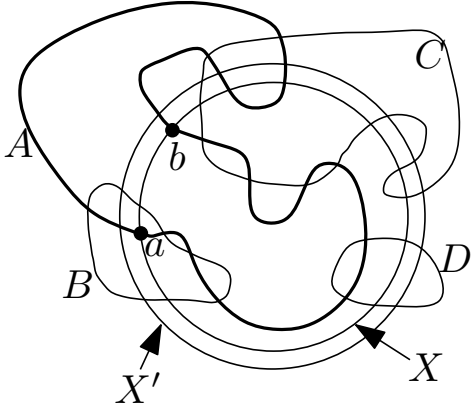


Figure 6: Before  $X$  pushes  $A$  and  $B$

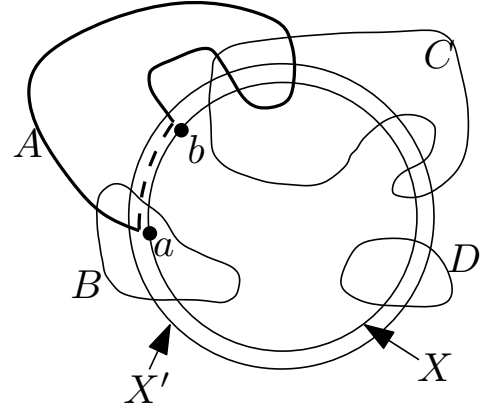


Figure 7: After  $X$  pushes  $A$  and  $B$

vertices as the number of intersections between  $\partial X$  and  $\partial R$ . As before, these vertices correspond to the intersections but are slightly perturbed from the intersections because of the pushing with non-zero gaps.

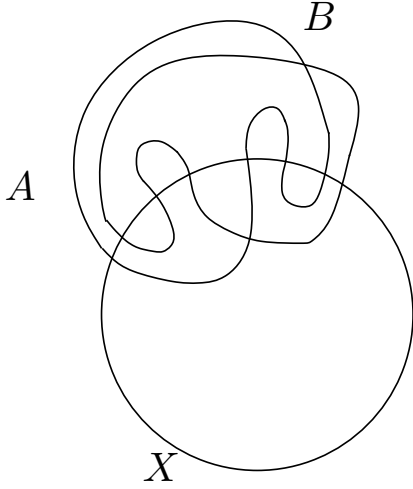


Figure 8: Before  $X$  pushes  $A$  and  $B$

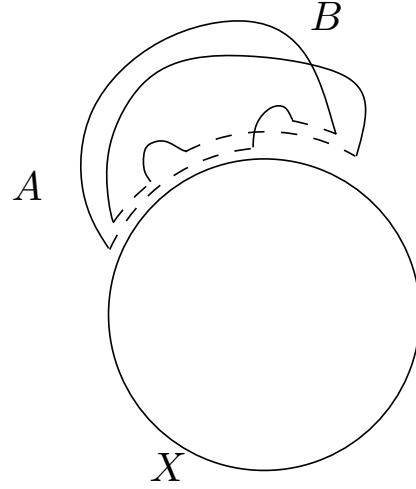


Figure 9: After  $X$  pushes  $A$  and  $B$

## C Lower-bounds

In this section we give examples of regions of low (but superlinear) union complexity that do not admit a PTAS.

**Observation C.1.** *The problem of approximating minimum-size set-cover is:*

1. APX-hard for a set  $\mathcal{R}$  of  $n$  4-sided polygons in the plane of union complexity  $n\alpha(n)$ .
2. inapproximable within  $o(\log s)$  factor for a set  $\mathcal{R}$  of  $n$   $4s$ -sided polygons in the plane of union complexity  $n2^{\alpha(n)^{O(s)}}$ , for any integer  $s > 3$ .
3. inapproximable within  $o(\log d)$  factor for a set  $\mathcal{R}$  of  $n$  halfspaces in  $\mathbb{R}^d$ , for any integer  $d > 3$ .

*Proof.* 1. Chan-Grant [8] showed that computing minimum size set-covers for objects defined by shadows of line-segments in the plane is APX-hard. By Davenport-Schinzel sequences, the union complexity of  $n$  line-segments in the plane is  $O(n\alpha(n))$ . These shadows can be ‘closed off’ without any further intersections to derive the 4-sided polygons.

2. Trevisan [32] showed that computing minimum size set-covers for general set systems  $(V, \{S_1, \dots, S_m\})$  where each  $S_i$  has size  $O(s)$  is inapproximable with factor  $o(\log s)$  unless  $P = NP$ . These sets can be easily implemented using regions whose boundaries cross at most  $O(s)$  times. To see this place a point corresponding to each vertex on the  $x$ -axis. Then for each set  $S_i$  construct  $x$ -monotone curve  $\gamma_i$  with  $2s + 1$  horizontal segments and  $2s$  nearly vertical segments so that the points corresponding to the vertices in  $S_i$  lie above  $\gamma_i$  and all other points lie below  $\gamma_i$ . These curves can easily be drawn in such a way that any two of them intersect at most  $O(s)$  times. Thus by bounds known on Davenport-Schinzel sequences, the lower envelope of the curves has complexity  $O(n2^{\alpha(n)^{O(s)}})$ . Thus if we consider the regions  $R_i$  defined by the set of points above  $\gamma_i$  we get a set of regions with small union complexity. These regions can be made bounded without increasing the union complexity.

3. There exist a set of points in  $\mathbb{R}^d$  (points on the moment curve; see Matousek [24]) such that every  $d/2$ -sized subset can be obtained by intersection with a halfspace. Thus a set-cover problem where every set has

size at most  $d/2$  can be realized with halfspaces in  $\mathbb{R}^d$ , which together with the Trevisan bound [32] implies the lower-bound.  $\square$

## D Separators. Proof of Theorem 2.2 and Theorem 2.3

In this section we prove Theorem 2.2 and Theorem 2.3. The proof of both these statements follow from a suitable subdivision of the plane, and the application of a variant of the planar graph separator theorem. Our proof can also be seen as a generalization of the separator theorem of Fox-Pach [18] where, given a set of curves with  $m$  intersections, they show the existence of a separator that intersects  $O(\sqrt{m})$  curves: this is obtained by applying the planar graph separator theorem on the arrangement induced by these curves (where each intersection is taken as a vertex). We also apply the planar graph separator theorem, but instead on a coarser subdivision of the plane. This subdivision is similar to a structure for the case of lines in the plane, called *cuttings* [23].

**Lemma D.1.** *Given a set  $S$  of  $n$   $x$ -monotone curves in the plane with  $m$  intersections (and where every pair of curves intersect  $O(1)$  times), and a parameter  $r$ , there exists a partition of  $\mathbb{R}^2$  into  $O(r + \frac{mr^2}{n^2})$  regions (each of constant descriptive complexity, defined by a constant number of curves of  $\Gamma(S)$  together with constant number of vertical line-segments) such that the interior of any region in this partition intersects curves of total weight  $O(n/r)$ .*

**Proof sketch.** We present the proof of the above lemma in Appendix E. Here we observe that a near-optimal (within log factors) result follows immediately from  $\epsilon$ -nets (the proof in Appendix A gets rid of these log factors using standard techniques from the theory of  $\epsilon$ -nets). For the purpose of designing QPTAS, however, it is not necessary to get rid of the log factors and the near-optimal bounds suffice at the expense of a slightly higher, but still quasi-polynomial, running time.

Given  $S$ , consider the set-system  $(S, \mathcal{F})$  induced by intersection with segments in the plane, i.e.,  $F \in \mathcal{F}$  iff there exists a line segment  $l$  s.t.  $F = \{s \in S \mid s \cap l \neq \emptyset\}$ . This set system can be shown to have a finite VC dimension [4]. Pick a random set  $R$  by uniformly adding each curve of  $S$  with probability  $p = (Cr \log r)/n$ , where  $C$  is a large constant. Then  $R$  is a  $(1/r)$ -net for  $(S, \mathcal{F})$  with probability at least  $9/10$  [24]. The expected size of  $R$  is  $np$ , and the expected number of intersections of curves in  $R$  is  $mp^2$ . By Markov's inequality, with probability at least  $9/10$ , the size of  $R$  is at most  $10np$ , and the number of intersections in  $R$  is at most  $10mp^2$ . Therefore with probability at least  $8/10$ ,  $R$  is a  $(1/r)$ -net and the size of the trapezoidal decomposition of  $R$  is  $O(r \log r + (mr^2 \log^2 r)/n^2)$ . Note that any open line-segment  $l$  in this trapezoidal decomposition must intersect at most  $n/r$  segments of  $S$ , as otherwise the set of curves intersecting  $l$  would not be hit by a curve from  $R$ , contradicting the fact that  $R$  is a  $(1/r)$ -net.

### Proof of Theorem 2.2

Let  $\mathcal{R} = \{R_1, \dots, R_n\}$  be the set of  $n$  weighted  $\alpha$ -simple regions, where weight of  $R_i$  is  $w_i$ , and  $R_i$  can be decomposed into  $v_i = v(R_i) \leq \alpha$   $x$ -monotone curves. In the standard way, by scaling, one can assume the weights are integral. The total weight is  $W = \sum_i w_i$ . First decompose each region  $R_i$  into  $v_i$   $x$ -monotone curves of weight  $w_i$  (and let  $S$  be the set of all such curves), and then each of these  $v_i$  curves for region  $R_i$  is replaced by  $\alpha w_i / v_i$  copies of weight 1 to get the set  $S'$ . Note that the total number of curves in  $S'$  from region  $R_i$  is  $\alpha w_i$ , and  $|S'| = \sum_i \alpha w_i = \alpha W$ . Apply Lemma D.1 to  $S'$  with the parameter  $r$  to be set later; if the curves in  $S'$  are disjoint, the proof of Lemma D.1 shows the existence of a subset  $R'$  of  $S'$  of total size  $O(r)$  such that the trapezoidal decomposition of  $R'$  gives a partition  $\mathcal{T}$  where the interior of any region in  $\mathcal{T}$  intersects  $O(\alpha W/r)$  curves of  $S'$ . Remove copies of the same curve in  $R'$  to get a subset  $R$  of  $S$  of total size  $O(r)$ . Now replace each  $s \in R$  with a small expanded copy of  $s$  which contains  $s$  in its interior and modify  $\mathcal{T}$  accordingly to use edges of this new region instead of  $s$ . As each curve  $s \in S$  of weight

$\Omega(\alpha W/r)$  must be present in  $R$  (otherwise it would intersect the interior of some region of  $\mathcal{T}$  contradicting the partitioning property), this ensures that the new triangulation  $\mathcal{T}'$  has the properties that *i*) every  $s \in S$  with weight  $\Omega(\alpha W/r)$  lies in the interior of a single face of  $\mathcal{T}'$ , *ii*) the number of vertices of  $\mathcal{T}'$  is  $O(r)$ , and *iii*) each edge in  $\mathcal{T}'$  intersects curves in  $S$  of total weight  $O(\alpha W/r)$  (for small-enough replacing regions around each  $s \in R$ ).

$\mathcal{T}'$  can be seen as an embedding of an underlying planar graph  $G$ . Give weights to each face of  $\mathcal{T}'$ : if a curve  $s \in S$  intersects  $t$  faces of  $\mathcal{T}'$ , add weight  $w(s)/t$  to the weight of each of these  $t$  faces. A variant of the planar graph separator theorem [26] now implies the existence of a simple cycle  $\mathcal{C}$  in  $\mathcal{T}'$  of  $O(\sqrt{r})$  vertices such that faces completely inside (and outside) have total weight at most  $2\alpha W/3$ , and hence so do the curves of  $S$  inside (and outside)  $\mathcal{C}$ . Let  $\mathcal{R}_{int} \subset \mathcal{R}$  be set of regions completely inside  $\mathcal{C}$ . Then as  $R_i$  produced curves of total weight  $\alpha w_i$ , we have  $\sum_{R_i \in \mathcal{R}_{int}} \alpha w_i \leq 2\alpha W/3$ , implying that stated bound on the total weight of regions in  $\mathcal{R}$  completely inside  $\mathcal{C}$ ,  $\sum_{R_i \in \mathcal{R}_{int}} w_i \leq 2W/3$ . Finally, the weight of the regions of  $\mathcal{R}$  intersected by  $\mathcal{C}$  is at most  $O(\sqrt{r}) \cdot O(\alpha W/r) = O(\alpha W/\sqrt{r})$ . Setting  $r = \alpha^2/\delta^2$  concludes the upper-bound.

Note that  $\mathcal{C}$  consists of alternating pieces of curves of  $R$  and vertical line-segments from the trapezoidal decomposition of  $R$ . Each vertical line-segment in this decomposition is between an endpoint of a curve of  $S$  and another curve of  $S$ . Thus  $\mathcal{C}$  can be specified by giving a sequence of  $O(\alpha/\delta)$  curves of  $S$  together with specifying which endpoint is used for every set of consecutive curves in the sequence. These specifications require an additional  $O(\alpha/\delta)$  bits. The optimality of this statement can be seen by the following construction where  $\mathcal{R}$  consists of a set of disjoint line segments of weight 1. Take a regular polygon  $\mathcal{P}$  with  $c/\delta$  vertices, and place  $\delta n/c$  copies of  $\mathcal{P}$  concentrically, each shrunk slightly more than the previous one so that there are no intersections between any two copies. Note that one can choose the scaling factor small-enough such that any closed curve separating two different copies of  $\mathcal{P}$  must also have at least  $c/\delta$  vertices. Finally replace each polygon with  $c/\delta$  line segments corresponding to its sides (slightly perturbed so that they are disjoint). Take any balanced closed curve  $\mathcal{C}'$  in the plane. If it contains at least one copy of  $\mathcal{P}$  completely inside, and one copy completely outside, then by construction it has at least  $c/\delta$  vertices. Otherwise, say there is no copy of  $\mathcal{P}$  completely inside  $\mathcal{C}'$ . As  $\mathcal{C}'$  is balanced, it contains at least  $n/3$  curves inside or intersecting its boundary; these curves belong to at least  $(n/3)/(c/\delta) = \delta n/(3c)$  different copies of  $\mathcal{P}$ , and each of these copies must intersect  $\mathcal{C}'$  in at least one curve.

### Proof of Theorem 2.3

Given the set  $\mathcal{R}$  of  $n$   $\alpha$ -simple regions with  $m$  intersections, construct from it the set  $S$  of  $O(\alpha n)$   $x$ -monotone curves. Apply Lemma D.1 to  $S$  to get a partition  $\mathcal{T}$  of  $\mathbb{R}^2$  into  $O(r + mr^2/\alpha^2 n^2)$  regions.  $\mathcal{T}$  can be seen as an embedding of an underlying planar graph  $G$ . Give weights to each face of  $\mathcal{T}$ : if a region  $s \in \mathcal{R}$  intersects  $t$  faces of  $\mathcal{T}$ , add weight  $1/t$  to the weight of each of these  $t$  faces. Now from [26] we get a simple cycle  $\mathcal{C}$  in  $\mathcal{T}$  of  $O(\sqrt{r + mr^2/\alpha^2 n^2})$  vertices such that faces completely inside (and outside) have total weight at most  $2n/3$ , and hence so do the regions of  $\mathcal{R}$  inside (and outside)  $\mathcal{C}$ . The weight of the regions of  $\mathcal{R}$  intersected by  $\mathcal{C}$  is at most  $O(\sqrt{r + mr^2/\alpha^2 n^2}) \cdot O(\alpha n/r) = O(\sqrt{m + \alpha^2 n^2/r})$ .



## E Proof of Lemma D.1

*Proof.* We briefly now review the basic partitioning method of using *trapezoidal decompositions*. Given a set  $R \subseteq S$  of  $x$ -monotone curves, one can partition the space (say inside a large-enough rectangle containing all the curves of  $S$ ) as follows. For each endpoint of a curve in  $R$  or an intersection-point between curves in  $R$ , shoot a vertical ray upwards (and downwards) till it hits another curve (or the bounding rectangle). The union of all these vertical segments together with  $R$  partitions the bounding rectangle into a set of regions. A crucial fact is that each region  $\Delta$  in this partition is determined by a constant (2, 3 or 4) number of curves in  $R$ . Call such regions *trapezoidal regions* (or trapezoids for brevity), and the partition is called a *trapezoidal decomposition*<sup>7</sup>. Denote by  $\Xi(R)$  this set of trapezoidal regions in the trapezoidal decomposition of  $R$ . The size,  $|\Xi(R)|$ , of the trapezoidal decomposition of  $R$  is the number of trapezoids in  $\Xi(R)$ ; it is, within a constant-factor, equal to the total number of end- and intersection- points in  $R$ . A trapezoid present in the trapezoidal decomposition of any subset  $R$  of  $S$  is called a *canonical trapezoid*. For a canonical trapezoid  $\Delta$ , let  $|\Delta|$  denote the set of curves of  $S$  intersected by  $\Delta$ . A trapezoid  $\Delta$  is present in the trapezoidal decomposition of  $R$  if and only if its determining curves are present in  $R$ , and  $R$  does not contain any of the curves of  $S$  that intersect  $\Delta$ . For the rest of the proof, we only work with canonical trapezoids determined by 4 curves. The case for canonical trapezoids determined by 2 and 3 curves is similar.

First note that a slightly weaker bound (within logarithmic factors) follows immediately from  $\epsilon$ -nets. Given  $S$ , consider the set-system  $(S, \mathcal{F})$  induced by intersection with segments in the plane, i.e.,

$$F \in \mathcal{F} \text{ iff there exists a line segment } l \text{ s.t. } F = \{s \in S \mid s \cap l \neq \emptyset\}$$

Pick a random set  $R$  by uniformly adding each curve of  $S$  with probability  $p = (Cr \log r)/n$ , where  $C$  is a large constant. Then  $R$  is a  $(1/r)$ -net for  $(S, \mathcal{F})$  with probability at least  $9/10$  [24]. The expected size of  $R$  is  $np$ , and the expected number of intersections of curves in  $R$  is  $mp^2$ . By Markov's inequality, with probability at least  $9/10$ , the size of  $R$  is at most  $10np$ , and the number of intersections in  $R$  is at most  $10mp^2$ . Therefore with probability at least  $8/10$ ,  $R$  is a  $(1/r)$ -net *and* the size of the trapezoidal decomposition of  $R$  is  $O(r \log r + (mr^2 \log^2 r)/n^2)$ . Note that any open line-segment  $l$  in this trapezoidal decomposition must intersect at most  $n/r$  segments of  $S$ , as otherwise the set of curves intersecting  $l$  would not be hit by a curve from  $R$ , contradicting the fact that  $R$  is a  $(1/r)$ -net.

Set  $p = Cr/n$  (for a small-enough constant  $C$  to be set later), and pick each curve in  $S$  with probability  $p$  to get a random sample  $R$ . Construct the trapezoidal decomposition  $\Xi(R)$  of  $R$ . If all trapezoids  $\Delta \in \Xi(R)$  intersect at most  $n/r$  curves in  $S$ , we are done. Otherwise we will further partition each violating  $\Delta$ , based on two ideas. First, the expected number of trapezoids in  $\Xi(R)$  intersecting more than  $n/r$  curves are few. In particular, we will show (Lemma E.2) that the expected number of trapezoids intersecting at least  $tn/r$  curves in  $S$  is *exponentially* decreasing as a function of  $t$ . Second, consider a  $\Delta$  intersecting a set, say  $S_\Delta$ , of  $n_\Delta = tn/r$  curves of  $S$ . Use the weaker bound on  $S_\Delta$  with parameter  $t$  to get a partition inside  $\Delta$  of  $O(t \log t + (m_\Delta t^2 \log^2 t)/n_\Delta^2) = O(t^2 \log^2 t)$  trapezoids. By definition, each such trapezoids intersects at most  $n_\Delta/t = n/r$  curves of  $S_\Delta$  (and hence of  $S$ ). Thus refining each  $\Delta$  gives the required partition on  $S$  with parameter  $r$ . It remains to bound the overall expected size of this partition.

**Lemma E.1.** *Given a set  $S$  of  $n$   $x$ -monotone curves in the plane with  $m$  intersections, the number of canonical trapezoids defined by  $S$  that intersect at most  $k$  curves of  $S$  is  $O(nk^3 + mk^2)$ .*

*Proof.* Let  $\Xi_{\leq k}$  be the set of canonical trapezoids defined by  $S$  that intersect at most  $k$  curves of  $S$ . The proof is standard via the Clarkson-Shor technique. Construct a sample  $T$  by adding each curve of  $S$  with

<sup>7</sup>We refer the reader to [24] for a nice exposition on trapezoidal decompositions.

probability  $p_0$ ; the expected total number of picked curves is  $np_0$  and the expected number of intersections between the curves of  $T$  is  $mp_0^2$ . The trick is to count the expected size of  $\Xi(T)$  in two ways. On one hand, it is at most  $O(np_0 + mp_0^2)$  (i.e., the expected number of vertices present in  $\Xi(T)$ ). On the other hand, as the probability of a canonical trapezoid  $\Delta$  being in  $\Xi(T)$  is  $p_0^4(1 - p_0)^{|\Delta \cap S|}$ , it is at least

$$\sum_{\Delta \in \Xi_{\leq k}} p_0^4(1 - p_0)^{|\Delta \cap S|} \geq \sum_{\Delta \in \Xi_{\leq k}} p_0^4(1 - p_0)^k$$

where the sum is over all canonical trapezoids  $\Delta$  which intersect at most  $k$  curves of  $S$ . Therefore,

$$\begin{aligned} \sum_{\Delta} p_0^4(1 - p_0)^k &= |\Xi_{\leq k}| \cdot p_0^4(1 - p_0)^k \leq E[|\Xi(T)|] = np_0 + mp_0^2 \\ |\Xi_{\leq k}| &\leq \frac{np_0 + mp_0^2}{p_0^4(1 - p_0)^k} = O(nk^3 + mk^2) \end{aligned}$$

for  $p_0 = 1/2k$ . □

**Lemma E.2.** *Expected number of trapezoids in  $\Xi(R)$  intersecting at least  $tn/r$  curves of  $S$  is*

$$O\left((t^3r + \frac{mr^2t^2}{n^2})e^{-t}\right)$$

*Proof.* By definition:

$$E[|\Delta \in \Xi(R) \text{ s.t. } |\Delta \cap S| = tn/r|] = |\Delta \text{ s.t. } |\Delta \cap S| = tn/r| \cdot p^4(1 - p)^{tn/r}$$

Using Lemma E.1,

$$\begin{aligned} E[|\Delta \in \Xi(R) \text{ s.t. } |\Delta \cap S| = tn/r|] &\leq O\left(n(tn/r)^3 + m(tn/r)^2\right) p^4(1 - p)^{tn/r} \\ &= O\left((t^3r + \frac{mr^2t^2}{n^2})e^{-t}\right) \end{aligned}$$

The bound follows by summing up over all trapezoids intersecting at least  $tn/r$  curves in  $S$ . □

Now we can complete the proof of the theorem. Let  $n_{\Delta} = t_{\Delta}n/r$  be the number of curves in  $S$  intersected by each trapezoid  $\Delta \in \Xi(R)$  (and  $m_{\Delta}$  the number of their intersections). Using the weaker bound, refine trapezoid  $\Delta$  by adding a  $(1/t_{\Delta})$ -net  $R_{\Delta}$  for all the  $t_{\Delta}n/r$  curves of  $S$  intersected by  $\Delta$ . The resulting

expected total size of the trapezoidal partition is:

$$\begin{aligned}
&= |R| + \sum_{\Delta} Pr[\Delta \in \Xi(R)] \cdot \text{Size of trapezoidal decomposition of } (1/t_{\Delta})\text{-net within } \Delta \\
&= |R| + \sum_{\Delta} Pr[\Delta \in \Xi(R)] \cdot O\left(t_{\Delta} \log t_{\Delta} + \frac{m_{\Delta} t_{\Delta}^2 \log^2 t_{\Delta}}{n_{\Delta}^2}\right) \text{ (using the weaker bound)} \\
&\leq |R| + \sum_{\Delta} Pr[\Delta \in \Xi(R)] \cdot O(t_{\Delta}^2 \log^2 t_{\Delta}) \text{ (as } m_{\Delta} \leq n_{\Delta}^2) \\
&= |R| + \sum_j \sum_{\substack{\Delta \text{ s.t.} \\ 2^j \leq t_{\Delta} \leq 2^{j+1}}} Pr[\Delta \in \Xi(R)] \cdot O(t_{\Delta}^2 \log^2 t_{\Delta}) \\
&\leq |R| + \sum_j E[\text{Number of trapezoids } \Delta \text{ in } \Xi(R) \text{ with } 2^j \leq t_{\Delta}] \cdot O(2^{2(j+1)} \log^2 2^{j+1}) \\
&\leq |R| + \sum_j O\left((2^{3j} r + \frac{mr^2 2^{2j}}{n^2}) e^{-2j}\right) \cdot O(2^{2(j+1)} \log^2 2^{j+1}) \text{ (Lemma E.2)} \\
&= |R| + r \sum_j O(2^{3j} e^{-2j}) \cdot O(2^{2(j+1)} \log^2 2^{j+1}) + \frac{mr^2}{n^2} \sum_j O(2^{2j} e^{-2j}) \cdot O(2^{2(j+1)} \log^2 2^{j+1}) \\
&= np + mp^2 + O(r) + O\left(\frac{mr^2}{n^2}\right) = O\left(r + \frac{mr^2}{n^2}\right) \text{ (the summands form a geometric series)}
\end{aligned}$$

as required. This finishes the proof of Lemma D.1. □